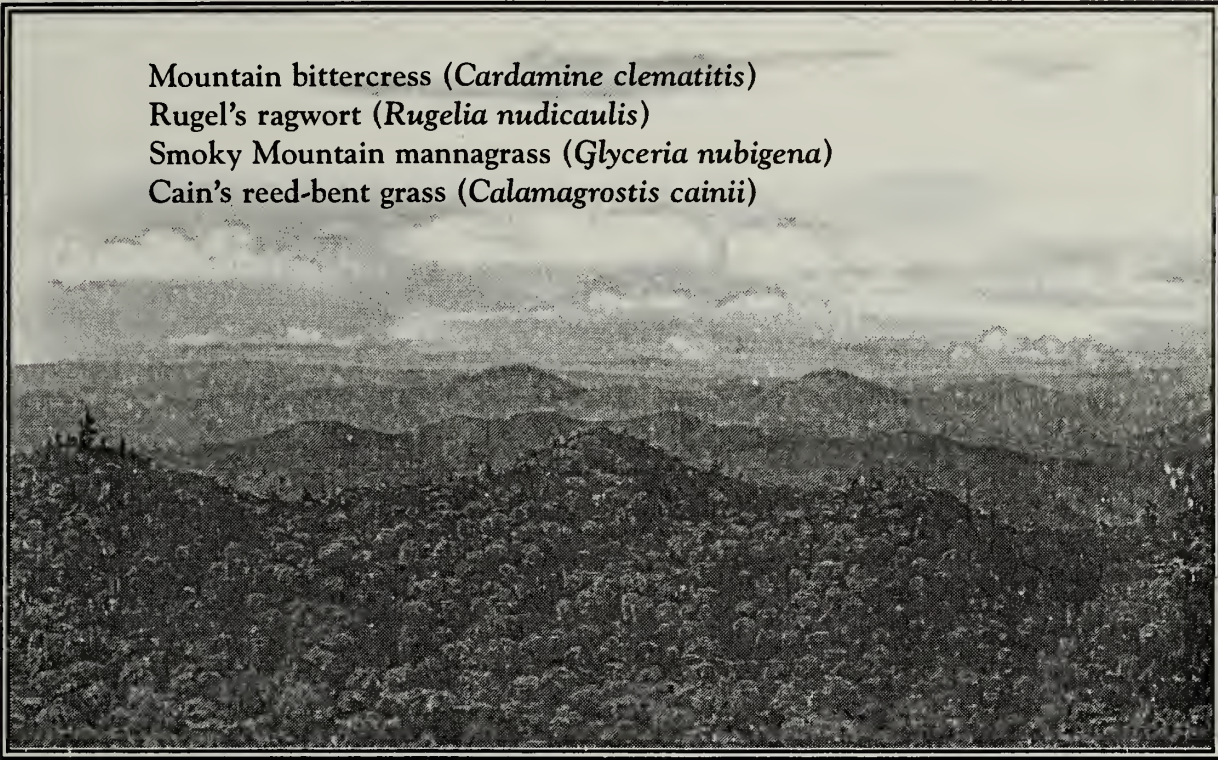


# Habitat Modeling and Conservation of Four Vascular Plants Endemic to the Southern Appalachian Mountains

Mountain bittercress (*Cardamine clematitis*)  
Rugel's ragwort (*Rugelia nudicaulis*)  
Smoky Mountain mannagrass (*Glyceria nubigena*)  
Cain's reed-bent grass (*Calamagrostis cainii*)




in Great Smoky  
Mountains  
National Park



National Park Service

107 Park Headquarters Road, Gatlinburg, Tennessee 37738

December, 1999



Digitized by the Internet Archive  
in 2012 with funding from  
LYRASIS Members and Sloan Foundation

<http://archive.org/details/habitatmodelingc00boet>

# Habitat Modeling and Conservation of Four Vascular Plants Endemic to the Southern Appalachian Mountains

- Mountain bittercress (*Cardamine clematitis*)
- Rugel's ragwort (*Rugelia nudicaulis*)
- Smoky Mountain mannagrass (*Glyceria nubigena*)
- Cain's reed bent-grass (*Calamagrostis cainii*)

John R. Boetsch  
University Of North Carolina

Janet Rock  
National Park Service

Great Smoky Mountains National Park  
107 Park Headquarters Road  
Gatlinburg, Tennessee 37738

December 1999

USGS-BRD Agreement No. 1445-IA09-96-0027

## CONFIDENTIALITY OF INFORMATION

“Information concerning the nature and specific location of a National Park System resource which is endangered, threatened, rare, or commercially valuable ... may be withheld from the public in response to a request under section 552 of Title 5 United States Code ....” Sec. 207. Confidentiality of Information. National Parks Omnibus Management Act of 1998.



## ACKNOWLEDGMENTS

This project was supported by the “Species at Risk” Program of the USGS Biological Resources Division (BRD) from 1996 through 1998. We would like to express our appreciation to Dr. Joe Clark and Dr. Frank van Manen of the Southern Appalachian Field Laboratory of USGS-BRD in Knoxville, Tennessee, for their indispensable assistance with the application of Mahalanobis distance to rare plant modeling and analyses of validation data. Thanks to Henry McNab (US Forest Service) for generously providing the code for Terrain Shape Index and Landform Index. We would also like to thank Dr. Peter White (University of North Carolina - Chapel Hill), Dr. John D. Peine (BRD), and Dr. Joe Clark for their valuable comments and suggestions on an earlier draft of the report. Many thanks to Stephanie Wilds for numerous and invaluable contributions to the project. We gratefully acknowledge those who provided location information and other data concerning the study species: Niki Nicholas (TVA), Claire Newell, Gary Kauffman (US Forest Service), Bob Dellinger (NPS), Mike Schafale (NC Heritage Program), Lisa Jameson (NPS), Nora Murdock (USFWS), and the Natural Heritage Program botanists from Tennessee, North Carolina, Georgia, South Carolina, and Virginia. We are indebted to the volunteers and technicians of Twin Creeks Natural Resource Center at GSMNP that so cheerfully endured long hours and wet boots: Eric Nielsen, Jennifer Dean, Aspen Madrone, Jillian Archer, Katy Crosthwaite, Erica Choberka, Jennifer Tietjen, and Brian Yahn. And, thanks to Keith Langdon for inspiration and Meryl Rose and Becky Nichols for excellent editorial comments. Finally, Steve Kemp, Great Smoky Mountains Natural History Association, generously designed and produced our report cover.





## ABSTRACT

We present the results of a two-year project evaluating the conservation status of four plant species endemic to the southern Appalachian Mountains of the southeastern United States. The project objectives were to identify the global distributions and habitat requirements of *Cardamine clematitis* (Mountain bittercress), *Rugelia nudicaulis* (Rugel's ragwort), *Glyceria nubigena* (Smoky Mountain mannagrass), and *Calamagrostis cainii* (Cain's reed bent-grass). All are perennial, herbaceous plants found at high elevations, primarily within Great Smoky Mountains National Park (GSMNP). Geographic Information System (GIS) models depicting the distribution of suitable habitat for each species within GSMNP were created during this study and are described in this report. One of these models was selected for field-testing and is the first time Mahalanobis distance ( $D^2$ ) has been used to model habitat for a rare plant species (*Cardamine clematitis*). Subsequent analysis determined that this model adequately defined distribution of suitable habitat. Habitat descriptions based on a combination of new and existing data, as well as summaries of known range distributions for each species, are included. All four species face potential threats due to ongoing changes in high elevation ecosystems throughout the region (e.g., acid precipitation, ground level ozone, etc.). Monitoring changes in population size and structure is needed to determine the response of these species to habitat change.

## EXECUTIVE SUMMARY

### *Cardamine clematitis*

*Cardamine clematitis* is presently known and documented from 18 counties in three states. At least 43 verified populations occur within GSMNP, with 27 additional populations outside the Park. Our modeling efforts reveal that habitat for this species is patchy and confined to high elevation drainages in central and eastern GSMNP. It is likely that additional populations will be located with further surveys.





### ***Rugelia nudicaulis***

*Rugelia nudicaulis* is only known to occur inside GSMNP. It is documented from two counties in North Carolina and two in Tennessee. Population boundaries are extremely difficult to assign to this species given its often continuous occurrence on mountain summits and side-slopes.

However, following current Natural Heritage Program criteria, there are 17 populations of *Rugelia* in the Park. Additional work on the genetic differentiation between populations and the response of *Rugelia* to changes in overstory composition (e.g., species, percent cover, etc.) is needed.

### ***Glyceria nubigena***

*Glyceria nubigena* is documented from two counties in Tennessee and three in North Carolina, with the majority of occurrences found in artificial clearings along trails and roadways where sunlight is typically more abundant. Estimation of numbers of populations based on distribution and boundaries is difficult to determine for this species, and its natural habitat requirements are not well understood. Further work is needed to establish criteria for distinguishing populations.

### ***Calamagrostis cainii***

*Calamagrostis cainii* is known from one county in Tennessee and two in North Carolina. It is restricted to only three mountain summits and is the rarest of the four species surveyed for this project. Only five populations are documented, three of which are found in GSMNP. The majority of total number of individuals is also found in GSMNP.



## Table of Contents

Acknowledgments .....	i
Abstract .....	ii
Executive Summary .....	ii
Introduction .....	1
Project Objectives .....	3
Methods and Materials .....	4
Data review .....	4
Location data .....	4
Habitat data .....	4
Preliminary field surveys .....	4
Location data .....	5
Habitat data .....	5
Habitat Model Development .....	6
Model Limitations .....	7
Data Analysis .....	9
Mahalanobis Distance .....	10
Field validation of predictive habitat models using <i>Cardamine</i> .....	10
Analysis of model output and field validation .....	12
Results and Discussion .....	15
Summary of known locations in GSMNP .....	15
<i>Cardamine clematitidis</i> .....	15
<i>Rugelia nudicaulis</i> .....	16
<i>Glyceria nubigena</i> .....	17
<i>Calamagrostis cainii</i> .....	17
Natural history, ecology, and habitat descriptions .....	17
<i>Cardamine clematitidis</i> .....	17
<i>Rugelia nudicaulis</i> .....	18
<i>Glyceria nubigena</i> .....	19
<i>Calamagrostis cainii</i> .....	20
Documented and potential threats .....	21
Genetic diversity .....	21
Air pollution .....	22
Introduced species .....	22
NPS activities .....	23
Data gaps .....	24
Model output and evaluation .....	25
Frequency of suitable habitat .....	31
<i>Cardamine clematitidis</i> field evaluation .....	32
<i>Cardamine clematitidis</i> suitable habitat .....	35
<i>Cardamine clematitidis</i> summary .....	36
Suggestions for model improvement .....	36



Conclusion and Recommendations .....	38
<i>Cardamine clematitidis</i> .....	38
<i>Rugelia nudicaulis</i> .....	38
<i>Glyceria nubigena</i> .....	39
<i>Calamagrostis cainii</i> .....	39
Summary .....	40
Literature Cited .....	41
Appendices .....	46
Appendix A: Herbaria visited during the course of this study .....	47
Appendix B: GIS variables used to model habitat in GSMNP .....	48





## List of Tables

Table 1:	Current status of study species .....	1
Table 2:	Weight factors used to improve representation of large populations of <i>Calamagrostis cainii</i> in location data set prior to modeling .....	8
Table 3:	Statistical summary for Mahalanobis distance ( $D^2$ ) output .....	25
Table 4:	Frequency of suitable habitat in GSMNP .....	31
Table 5:	Presence/absence of <i>Cardamine clematitis</i> in validation plots, GSMNP .....	33
Table 6:	Chi-square analysis of <i>Cardamine clematitis</i> presence in validation plots, as grouped by validation classes .....	33

## List of Figures

Figure 1:	Cumulative frequency of $D^2$ values for <i>Cardamine clematitis</i> positions .....	26
Figure 2:	Cumulative frequency of $D^2$ values for <i>Rugelia nudicaulis</i> positions .....	27
Figure 3:	Cumulative frequency of $D^2$ values for <i>Glyceria nubigena</i> positions .....	28
Figure 4:	Cumulative frequency of $D^2$ values for <i>Calamagrostis cainii</i> positions .....	29
Figure 5:	Classification of validation data for <i>Cardamine clematitis</i> .....	35



## List of Maps

- Map 1: Known locations of *Cardamine clematitis* in GSMNP
- Map 2: Known locations of *Rugelia nudicaulis* in GSMNP
- Map 3: Known locations of *Glyceria nubigena* in GSMNP
- Map 4: Known locations of *Calamagrostis cainii* in GSMNP
- Map 5: Predicted distribution of *Cardamine clematitis* in GSMNP
- Map 6: Predicted distribution of *Cardamine clematitis* in Clingmans  
Dome area, GSMNP
- Map 7: Predicted distribution of *Rugelia nudicaulis* in GSMNP
- Map 8: Predicted distribution of *Rugelia nudicaulis* in Clingmans  
Dome area, GSMNP
- Map 9: Predicted distribution of *Glyceria nubigena* in GSMNP
- Map 10: Predicted distribution of *Glyceria nubigena* in Clingmans  
Dome area, GSMNP
- Map 11: Predicted distribution of *Calamagrostis cainii* in GSMNP
- Map 12: Predicted distribution of *Calamagrostis cainii* in  
Mount Le Conte area, GSMNP
- Map 13: Predicted distribution of *Cardamine clematitis* in GSMNP,  
as defined by field validation
- Map 14: Validated distribution of *Cardamine clematitis* near Clingmans  
Dome, GSMNP



## INTRODUCTION

The central purpose of this study was to review the status of four plant species endemic to the southern Appalachian Mountains, and to produce predictive habitat models. The study species are as follows:

<u>Scientific name and authority</u>	<u>Common name</u>	<u>Family</u>
<i>Cardamine clematitis</i> Shuttleworth ex A. Gray	Mountain bittercress	Brassicaceae
<i>Rugelia nudicaulis</i> Shuttleworth ex Chapman	Rugel's ragwort	Asteraceae
<i>Glyceria nubigena</i> W. A. Anderson	Smoky Mountain mannagrass	Poaceae
<i>Calamagrostis cainii</i> A. S. Hitchcock	Cain's reed bent-grass	Poaceae

All four species have similar distributions that follow the extent of higher summits in the southern Appalachians and are restricted to elevations above approximately 1,000 meters (3,280 feet).

Three of the four are located only in the mountains of western North Carolina and eastern

Tennessee, whereas *Cardamine clematitis* is also located in extreme southwestern Virginia.

*Calamagrostis* is listed as endangered in the states of North Carolina and Tennessee; *Glyceria* and

*Rugelia* are each listed as threatened in both states; and *Cardamine* is listed as threatened in

Tennessee, but presently is not listed for Virginia (Table 1).

**Table 1:** Current status of study species.<sup>1</sup>

Species	Tennessee	North Carolina	Virginia	Federal	TNC
<i>Cardamine clematitis</i>	S2 – Threatened	S2? - Candidate	S1S2 - not listed	Species of Concern (C2)	G2G3
<i>Rugelia nudicaulis</i>	S2 – Endangered	S3 - Threatened	not present	Species of Concern (C2)	G3
<i>Glyceria nubigena</i>	S2 – Threatened	S2 - Threatened	not present	Species of Concern (C2)	G2
<i>Calamagrostis cainii</i>	S1 – Endangered	S1 - Endangered	not present	Species of Concern (C2)	G1

<sup>1</sup> The Nature Conservancy ranks species on a “global” (G) level based on the number of known populations for that species plus other factors that may be contributing to a documented or predicted decline in abundance. Ranks are as follows: G1 = critically imperiled globally due to





extreme rarity, with five or fewer populations or very few remaining individuals; G2 = critically imperiled with six to 20 populations or few remaining individuals; G3 = very rare and local throughout its range or found locally within a restricted range; G4 = apparently secure globally though it may be rare in parts of its range; and G5 = demonstrably secure globally though it may be rare in parts of its range. State rankings (S1 to S5) follow the same system as global ranks, but are determined by Natural Heritage Programs based within each state (Amoroso 1997, Nordman 1996, and Ludwig 1997).

Selection of these four species was based on the following criteria:

- (1) All four are narrow endemics of uncertain conservation status and believed to be rare;
- (2) A majority of known occurrences for each species is found within Great Smoky Mountains National Park (GSMNP), a 200,000-hectare International Biosphere Reserve that straddles the boundary between Tennessee and North Carolina;
- (3) The U.S. Fish and Wildlife Service (USFWS) previously listed each species as a Category 2 species (Species of Concern); USFWS discontinued listing species in this category in 1997. Category 2 designation formerly meant that current information "...indicates that proposing to list them as endangered or threatened species is possibly appropriate but for which substantial data on biological vulnerability and threats are not currently known or on file to support the immediate preparation of rules." (*Federal Register*, September 27, 1985);
- (4) All four occur at high elevations in the southern Appalachians, and may be affected by current environmental and ecological changes (e.g., air quality, non-native pests, etc.) in the communities in which they are found. The nature of these changes and their effect on the study species are not well understood at present.

In this report we explain the methods used to describe habitat and build predictive models. We then summarize known occurrence and habitat requirements for each species, identify potential threats, and describe results of predictive habitat modeling. We discuss the significance of this information in light of the current federal status of each species. Finally, we identify future conservation needs and propose additional avenues of research.



## **Project objectives**

The project objectives were to:

- (1) summarize the global distribution of each of the four study species, and
- (2) identify suitable habitat for each species within GSMNP.

The rationale for gathering this information is as follows: by knowing the extent and frequency of occurrence of each species – at both local and regional scales – better decisions can be made about present and future conservation needs. Evaluation of habitat requirements also advances further study of potential threats to the long-term viability of each species.

GSMNP is a 200,000-hectare (500,000-acre) protected area in Tennessee and North Carolina that contains approximately 75% of the remaining Fraser fir/red spruce forests in the world (source: USDA Forest Service aerial photography data, from 1985). Approximately 20% of GSMNP has never been logged, with much of this old-growth concentrated in the middle-to-high elevations of central and eastern GSMNP (Pyle 1985; Pyle 1988). A majority of the occurrences for all four species are found within GSMNP. By centering our efforts on GSMNP, we are taking advantage of the fact that this relatively large and intact reserve is already under federal protection.

By building predictive habitat distribution models, we can address the following questions:

- (1) What is the distribution and abundance of suitable habitat for each study species in GSMNP?
- (2) How fully do these rare species occupy or “saturate” their potential habitat?
- (3) Specifically, can 30-meter digital GIS data sufficiently delineate habitat?



## METHODS and MATERIALS

### Data review

#### Location data

Location data on *Cardamine*, *Calamagrostis*, *Glyceria*, and *Rugelia* were gathered from the Biological Conservation Data System (BCD) maintained by GSMNP, and state Natural Heritage Programs of North Carolina, Tennessee, and Virginia. Additional sources of information included journal articles, direct communication with regional botanists, and herbarium records (Appendix A). Several new occurrences, as well as errors and omissions, were identified during the process of compiling, reviewing, and verifying existing data.

#### Habitat data

Sources for habitat data included Tennessee Valley Authority (TVA), North Carolina Vegetation Survey (NCVS), and GSMNP long-term monitoring data from the 1970s and 1980s. Additional data were obtained from GSMNP's Rare Plant Monitoring Program. This program has monitored two populations of *Rugelia* (since 1989) and one population of *Cardamine* (since 1992).

Two different approaches to habitat characterization were used. They are as follows:

- (1) identifying plant species and small-scale environmental characteristics associated with the occurrence of each species to recognize suitable habitat in the field;
- (2) developing predictive models depicting the geographic extent and frequency of suitable habitat across the Park.

### Preliminary field surveys

Field surveys were conducted to relocate and confirm all known populations of all four species within GSMNP. Limited reconnaissance of populations occurring outside GSMNP was conducted primarily to verify occurrence and to identify obvious differences in habitat conditions





outside GSMNP. All known occurrences (inside and outside GSMNP) for *Calamagrostis* and *Rugelia* were visited. All *Glyceria* and *Cardamine* occurrences within GSMNP and approximately half of the known occurrences outside GSMNP also were visited during the course of this project.

### Location data

Our goal was to verify reported locations and improve location precision using a global positioning system (GPS) unit. The impetus for this work was to build an accurate set of location coordinates for which geographic data could be collected within GIS. This geographic data set would then serve as the basis for habitat modeling. Location verification and coordinate precision was important because existing information was frequently out-dated, in error, or of insufficient resolution to build a model. Location descriptions such as “mountains near Gatlinburg” or “Clingmans Dome area” are too vague for the purposes of mapping and sampling GIS data. Since the resolution of the Park’s GIS database is 30 meters, we attempted to match that resolution in the location data set. Whenever satellite reception was of sufficient quality, positions were recorded with an unassisted, military Y-code signal from the Department of Defense’s GPS (PLGR+96, Rockwell International, Cedar Rapids, Iowa). Use of a GPS unit enabled accurate positioning and navigation in rugged terrain away from trails. The unit typically yielded estimated horizontal displacement errors between 10 and 20 meters. When satellite reception was hindered by dense vegetation or steep terrain, positions were hand-mapped by relating altimeter and compass readings to local features on USGS topographic maps (7.5 minute; 1:24,000). Universal Transverse Mercator (UTM) coordinates were derived from positions mapped with an acetate overlay.

### Habitat data

Habitat data collected at field locations typically included site characteristics such as slope, aspect, elevation, and the aspect-dominant species in the canopy, understory, and herbaceous layers. Occasionally, more detailed data were collected on percent cover of associated plant



species, total canopy cover, and existing or potential threats (e.g., non-native insect infestations, wild boar rooting, etc.). Population data included estimates of population size, health (including evidence of reproduction, disease, and herbivory), and area covered. Voucher specimens were collected to document species occurrences. These specimens were deposited at GSMNP and University of North Carolina herbaria.

### **Habitat model development**

We limited our modeling efforts to GSMNP because the Park comprises the majority of known occurrences for each study species and is believed to provide the greatest amount of additional occurrences based on potential habitat. The availability of a geographically-referenced database for GSMNP permitted the spatial analysis of species distributions and the development of predictive habitat distribution models. GSMNP maintains numerous themes and calculated indices in GIS format. There are three types of data represented: raster grids, polygons, and points. Raster themes include forest type, geology, disturbance, elevation, and themes calculated from digital elevation: aspect, slope, curvature, Relative Slope Position (RSP), Topographic Relative Moisture Index (TRMI), Shannon-Weiner Index (SWI), and Topographic Convergence Index (TCI). Polygon themes include trails, streams, roads, watershed boundaries, the Park boundary, and the borders for the USGS topographic maps. Vegetation plot positions and the locations of distinct features such as home sites and exotic plant locations are examples of point themes.

In principle, habitat modeling uses data corresponding to known locations for a species to estimate the likelihood of a species' presence or absence given a new combination of input data. Such a model becomes potentially valuable when that estimation is automated for large geographic areas. We used the Park's existing GIS data as the basis for habitat modeling because the Park covers a large geographic area. By selecting GIS data as the foundation of our habitat models, we were able to produce models that depict the spatial extent, frequency, and distribution



of suitable habitat for each species – none of which would have been possible had we selected field data as the basis for modeling.

### Model Limitations

Despite these advantages, we were still confined by the limits of a GIS data set that lacks data on critical habitat factors such as soil type and vegetation structure. Also, the database has a maximum resolution of 30 meters, and, therefore, cannot represent microhabitat factors that function below that scale.

We used 14 GIS raster variables to characterize habitat conditions for input locations (Appendix B). The list of variables includes: (1) continuous variables such as elevation, slope, aspect (transformed to a continuous variable); (2) scalar variables such as relative slope position, Shannon-Weiner index, and Topographic Convergence index; and (3) categorical variables such as vegetation and rock outcrops. All variables in Appendix B were derived from the digital elevation model (DEM) except for vegetation and rock outcrop themes. The rock outcrop theme was created during the course of this project by digitizing these features from Park aerial photographs. This theme was included in the model for *Calamagrostis* because high elevation rock outcrops and landslide scars have been identified as critical habitat for this species (Wiser 1991; Wiser et al. 1996). Data on geochemistry were not included in the model construction because there was only very limited information available (e.g., soil type) and because coverage was not complete for GSMNP (e.g., parent geology).

Only locations away from trails with relatively high precision (GPS-documented below 20 meters) were used to sample the GIS for input data. The distribution models were based on the following number of unique coordinate positions: *Cardamine* - 187 positions, *Rugelia* - 124 positions, *Glyceria* - 40 positions, and *Calamagrostis* - 25 positions. It is important to note that the number of coordinate positions is not the same as the number of populations. Rather, several distinct coordinate positions may describe a single population, each with a unique combination of GIS data values.





Trailside occurrences were excluded from the input data set despite the fact that they comprised a large proportion of the total occurrences for *Glyceria* and *Rugelia*. This was done in order to remove the confounding factor of “noise” in these species’ response to meso-scale environmental variables presented by trailside occurrences. For example, along trails, herbaceous understory species are likely to respond much more to highly localized light availability and small-scale disturbance than to larger-scale variables such as elevation, aspect, and landscape protection. Since our models were based on GIS variables which had no way of incorporating such small-scale information, we decided it was best to remove trailside occurrences from our input data sets in the interest of better determining distributions away from trails.

Abundance data collected during preliminary field surveys indicated that population size for *Calamagrostis* was markedly heterogeneous, with a large proportion of the total number of stems of this species found in only a few populations. For this reason, abundance data for each *Calamagrostis* occurrence was used to weight the representation of the position data in the model building process. Because the cover of *Calamagrostis* had such a large range between sites (i.e., from less than 0.5 m<sup>2</sup> to 600 m<sup>2</sup>), failing to apply any weight factor would have over-represented the influence of small populations. Cover was weighted for this model according to the following table:

**Table 2:** Weight factors used to improve representation of large populations of *Calamagrostis* in location data set prior to modeling.

Cover less than/equal to:	Weight factor:	Frequency:
10 m <sup>2</sup>	1	14
20 m <sup>2</sup>	2	4
40 m <sup>2</sup>	3	2
80 m <sup>2</sup>	4	3
160 m <sup>2</sup>	5	0
320 m <sup>2</sup>	6	1
640 m <sup>2</sup>	7	1



Fourteen input locations had a weight factor of one, and only one had a weight factor of seven. By replicating each occurrence in the data set according to its weight factor, the *Calamagrostis* input locations were expanded from 25 positions to 53. This procedure was only applied to *Calamagrostis* because abundance data were much more consistent between sites for *Cardamine*, *Rugelia*, and *Glyceria*.

Since the vegetation theme is categorical rather than continuous, it was necessary to use design variables by creating a data layer for each cover type, with a value of 1 representing that type and a value of 0 representing all other cover types. The number of design variables was selected independently for each species, depending on the distribution of input locations across the vegetation types. We created four design variables for *Cardamine*. Specifically, because over 95% of the positions fell into four forest types (i.e., spruce-fir, northern hardwood, cove hardwood, and mesic oak hardwood), we created separate layers for each. Similarly, we created two design variables for *Rugelia* and *Glyceria* because approximately 88% of occurrences were found in spruce-fir and northern hardwood forests. We created a single design variable for *Calamagrostis* because over 86% of the input positions (weighted) was found in spruce-fir forest.

### **Data Analysis**

We selected multivariate analysis for this study based on the assumption that plants are more likely to be affected by complex combinations of many different habitat factors. Habitat models can be built from input data in two separate ways.

The first, called Boolean or logical analysis, defines habitat by the range of values that match the species' response for each of the selected variables. Areas falling within the range of values for the input data constitute suitable habitat that is defined by the overlap of suitable areas for each variable. Areas that fall outside the range of values for the input data for one or more variables are considered unsuitable. Despite its simplicity, the disadvantages of this technique include the following: (1) the importance of each variable included in the analysis tends to be overemphasized; (2) a few atypical occurrences can greatly expand the envelope of potential



habitat; and (3) this technique ignores the interactions between multiple variables and their influence on the response of the species.

The alternative to Boolean analysis comprises a broad array of modeling techniques that use multivariate statistics to estimate the contribution of each variable. While these models are more complex, they are more appropriate because they allow multiple, unique combinations of variables to be considered suitable habitat. In addition, multivariate models regard habitat as a continuum from suitable to unsuitable.

### Mahalanobis Distance

Mahalanobis distance ( $D^2$ ), the multivariate modeling technique used in this study, is a unitless, multivariate statistic that describes the deviation from an optimal set of conditions as defined by those inputs used to generate the model (Rao 1952; Clark et al. 1993). Higher values represent a set of conditions dissimilar from the optimum, whereas low values represent conditions similar to those supporting the model species. We calculated components of  $D^2$  in SAS (SAS Institute, Inc., 1990) based on habitat data from input locations using the following equation:

$$\text{Mahalanobis distance} = (x - \hat{\mu})' \Sigma^{-1} (x - \hat{\mu}),$$

where  $x$  is a vector of habitat characteristics in the GIS,  $\hat{\mu}$  is the mean vector of habitat characteristics of the input data, and  $\Sigma^{-1}$  is the inverse of the variance-covariance matrix calculated from the input data. The components were assembled to create an Arc/Info ® (ESRI, Redlands, California) macro language file that calculated  $D^2$  for each 30 x 30 - meter pixel in our study area.

### **Field validation of predictive habitat models using *Cardamine***

Modeling efforts usually do not include field validation of their predictive models because this requires a lot of time and resources. We decided it prudent to evaluate the accuracy of the modeling process and the assumptions made during model development. *Cardamine* was selected for field validation because the model for this species was based on a relatively large number of





initial positions. Given limitations in time and resources, we believed it preferable to more thoroughly understand the capabilities and limitations of one habitat model rather than have only limited analyses for all four models.

We defined our validation area as the portions of GSMNP above 914 meters elevation (3,000 feet). The validation area included 121,081 hectares comprising 59% of GSMNP. Elevations in the study area ranged from 914 to 2,025 meters (3,000 to 6,643 feet). This eliminated most lower-elevation pine and pine-oak communities from consideration, encompassed known occurrences of the study species in GSMNP to date, and allowed concentrated field effort in areas more likely to be occupied by *Cardamine*. Our definition of the study area tested the model at a narrow range of habitat conditions, and, as such, provided a conservative evaluation of the model.

Validation protocol was designed to evaluate whether the occurrence of *Cardamine* is inversely correlated with increasing values of  $D^2$ . Because we wanted to evaluate the model's capability for predicting absence as well as presence, we devised a validation scheme that sampled field positions across the entire range of  $D^2$  values. However, due to the rarity of the species, sampling the entire range of these values would result in a disproportionate number of survey plots where the species is absent compared with plots where the species is present. We defined six classes according to the amount of area they represented in  $D^2$  coverage by using the following formula:

$$\text{Pixels in validation area} = n + 2n + 4n + 8n + 16n + 32n$$

Thus  $n$  represents 1.587% of the pixels. We used this method to determine at which  $D^2$  value each class division should occur:

class 1: 3.748 - 10.51405

class 2:  $\leq 13.36148$

class 3:  $\leq 17.29093$

class 4:  $\leq 23.98191$

class 5:  $\leq 39.67574$

class 6:  $> 39.67574$  (maximum  $D^2$  for the study area was 1164.9)





We sampled each class of  $D^2$  values with 20 random validation plots. This method concentrated the field effort toward smaller values for  $D^2$ .

The validation area was confined to a zone between 100-500 meters (328-1,641 feet) from trails and roads. A minimum distance was established to reduce confounding effects, such as roads and trails, on plant distributions. A maximum distance also was set to limit the amount of time spent locating validation positions. Although this approach restricted validation to a subset (approximately one-third) of the study area, efficient sampling of a large number of validation points could not have been accomplished given the study area's extremely rugged terrain and large size. Proportion of pixels in each of the six classes was not different from that for the entire study area. The validation area was 42,236 hectares (469,291 pixels at 30-meter resolution, or 0.09 ha/pixel) which represented 34.8% of the study area and 20.5% of GSMNP. Coordinates for 120 validation plots (six in each class) were charted onto USGS 1:24,000 topographic maps.

Field validation took place during Fall 1997 and Spring 1998. Ninety percent of the validation plots ( $n = 108$ ) were located with a military Y-code GPS receiver (mean displacement error = 11.7 m, median = 9.5 m). The remaining 12 plots were located with map, compass, and altimeter. Once a sampling point was established, a 0.1-hectare plot (31.6 x 31.6 meters) centered on the point was surveyed. The presence or absence of *Cardamine* was recorded in each plot, along with additional ecological information characterizing the site.

### **Analysis of model output and field validation**

In the absence of field validation, model suitability can be evaluated by how well the model fits the data that went into its construction (i.e.,  $D^2$  values associated with model input locations). The power of a model can be determined by how well the model predicts independent occurrences that were either withheld during model development or were found subsequent to model development. For our analysis, we graphed cumulative frequency of  $D^2$  values associated with different sets of known occurrences. We then compared the shapes and placement along the



ordinate axis of each distribution curve (i.e., input locations versus withheld locations, etc.). The shape of each cumulative frequency curve indicates the tightness of the distribution; a shift to the left or right along the ordinate axis indicates a difference in the range of values.

*Cardamine* occurrences found after model development were compared with those used to build the model as well as with those found in the independent validation plots. For *Rugelia*, because numerous verified coordinate pairs were available, half of the positions were withheld from the model during development. Cumulative frequency curves for inputs and *Rugelia*'s positions were compared to see how well the model predicted independent occurrences. For both *Glyceria* and *Rugelia*, a large proportion of population occurrences was found along trails and roadways. The predictability of these occurrences was also evaluated with the cumulative frequency graphs.

The model output for each species was classified into three categories of habitat suitability. Although contrary to the purposes of describing habitat as a continuum from good to poor, by categorizing output we are able to generate maps that depict model output. Definition of each category was based on range of  $D^2$  values corresponding to input locations (i.e., positions to build models).  $D^2$  values were sorted in rank order, from low to high, since Mahalanobis distance is a dissimilarity index. The upper limits for the first two categories were assigned based on  $D^2$  values inclusive of the first 50% of input positions, and the first 95% of input positions. The last category, by default, included the tail 5% of input positions as well as presumably unsuitable habitat associated with higher  $D^2$  values. Assuming an inverse relationship between  $D^2$  and habitat quality, the first category represents optimal habitat conversely, the last category represents poor to unsuitable habitat. Habitat categories for each model were defined using the range of values for positions used to build the model.

Category 1 habitat (i.e., "optimal") was defined as the range of values inclusive of the first 50% of input occurrences.

Category 2 habitat (i.e., "moderately suitable") included the next 45% of inputs

Category 3 habitat (i.e., "poor to unsuitable") included all  $D^2$  values above the upper limit for Category 2.



*Cardamine* model field validation was analyzed using logistic regression (Hosmer and Lemeshow 1989) to determine the relationship between presence/absence of *Cardamine* and  $D^2$  values of corresponding pixels. We used chi-square analysis to test for differences between expected and observed distribution of *Cardamine* within stratified validation classes.





## RESULTS AND DISCUSSION

### Summary of known locations

The distribution of known locations is depicted for each species in Maps 1 through 4. Positions on these maps reflect a compilation of all known positions within GSMNP and include both trailside and off-trail positions. The following paragraphs summarize the number of occurrences and populations for each species. There is no known documentation of the pollen or seed dispersal mechanisms for these species and so our efforts at grouping occurrences into populations are necessarily suspect. However, we made every effort to consider the probable dispersal mechanisms for each species and to maintain consistency with The Nature Conservancy and state Natural Heritage Programs' protocols. Also, the distinction between occurrence and population should be clarified. For our purposes, we define an occurrence (synonymous with position and location in this document) as a set of coordinates with less than 50 meters estimated horizontal posting error; a population is an ecological entity that may be comprised of one to several occurrences.

### *Cardamine clematitis*

*Cardamine clematitis* has the broadest distribution of the four study species. It is found from Mt. Rogers Recreation Area in southwestern Virginia, south into Nantahala National Forest in western North Carolina. There have been reports of *Cardamine* in West Virginia (Strausbaugh and Core 1978), Georgia (Duncan and Kartesz 1981, Godfrey and Wooten 1981; Wofford 1989; Gleason and Cronquist 1991; Radford et al. 1968; Al-Shehbaz 1988), South Carolina (Gaddy et al. 1984; Wofford 1989; Gleason and Cronquist, 1991), and Alabama (Fernald 1950; Small 1933; Al-Shehbaz 1988; Rollins 1993). By examining specimens at several regional herbaria (see Appendix A), we determined that voucher specimens associated with these claims were actually misidentified specimens of a similar species, *Cardamine flagellifera* O.E. Schulz. We treated these two species as distinct (Dudley 1974; Al-Shehbaz 1988). *Cardamine* is presently known from the following 18 counties: Smyth, VA; Grayson, VA; Blount, TN; Carter, TN; Johnson, TN;



Sevier, TN; Unicoi, TN; Washington, TN; Avery, NC; Caldwell, NC; Graham, NC; Haywood, NC; Jackson, NC; Macon, NC; Mitchell, NC; Swain, NC; Watauga, NC; and Yancey, NC. At the start of this study, only 18 individual occurrences were documented for *Cardamine* within GSMNP. Our surveys within the Park group *Cardamine* into 18 populations (comprising a total of 266 occurrences) (Map 1). Outside GSMNP, the number of populations has tentatively been increased from 24 to 27. We believe this number would decrease if we applied TNC protocols for population delineation (i.e., occurrences within one kilometer of each other comprise a single population. Continued searches both within and outside GSMNP may reveal additional populations.

### *Rugelia nudicaulis*

*Rugelia nudicaulis* is endemic to GSMNP, although suitable habitat apparently exists in other spruce-fir areas in the southern Appalachians. It is known to occur from the Mt. Sterling area in eastern GSMNP to Hazel Creek drainage in central GSMNP (Map 2). It is presently documented from the following four counties: Cocke, TN; Sevier, TN; Haywood, NC; and Swain, NC. When considered in total, trailside and off-trail occurrences for *Rugelia* comprise approximately 12 large populations associated with the major mountain summits and slopes in eastern and central GSMNP. Our surveys located 248 off-trail occurrences and 531 trailside occurrences from an original documentation of only 45 occurrences. When considered independently, off-trail occurrences also comprise roughly 12 separate clusters with most of these occurring in close proximity (within 100 meters) to trailside populations (perhaps sub-populations). Further surveys would likely reveal plants dispersed between populations and clusters, resulting in a smaller overall number of populations. This species presented difficulties regarding field surveys simply because its populations cover so much area. Determination of numbers of populations is made more difficult because *Rugelia* is visited, and possibly pollinated, by wide-ranging bumblebees (*Bombus* spp.) and cuckoo bees (*Psithyrus* spp.) which might broaden the range of gene transfer within a population (J. Rock, unpublished data, GSMNP).



*Glyceria nubigena*

*Glyceria nubigena* is presently documented from five counties: Monroe, TN; Sevier, TN; Graham, NC; Haywood, NC; and Swain, NC. The majority of known occurrences are found along trails in GSMNP. Trailside occurrences are inherently difficult to group into functional populations. After data collection, we distinguished 13 populations (comprising 490 trailside occurrences and 40 off-trail occurrences) (Map 3). Further work should be devoted to establishing criteria for defining *Glyceria* populations given the contiguity of occurrences.

*Calamagrostis cainii*

*Calamagrostis cainii* is the most restricted in distribution and habitat specificity. *Calamagrostis* is only known from three distinct mountain summits, with the majority of individuals in one or two populations. Within GSMNP, this species is found only within a 6.4 kilometer radius of the summit of Mount Le Conte (Map 4). Five populations (comprising approximately 60 occurrences) have been documented – three of which are found in GSMNP. It is found in three counties: Sevier, TN; Buncombe, NC; and Yancey, NC.

**Natural history, ecology, and habitat descriptions**

Natural history observations, habitat data, and small-scale environmental characteristics, collected during field surveys were compiled to identify and refine suitable habitat.

*Cardamine clematitis*

*Cardamine clematitis* is a herbaceous perennial that blooms in May and fruits from May through July. This plant is clonal, with a shallow-rooted rhizome that bears one to several rosettes of waxy, evergreen, kidney-shaped basal leaves. Typically, a small fraction (less than 5%) of the total number of rosettes are fertile in any given population. We have no data on possible pollinators of *Cardamine*, but flower morphology indicates a generalist visitor. *Cardamine* has explosively-dehiscent fruits that may scatter seeds within a radius of a few meters. The seeds lack





any obvious dispersal adaptation, so the probable mechanism for dispersal is the explosive dehiscence combined with gravity, and possibly water, transport. Often, immature fruits were noted to have only a few developing seeds within. This raises questions as to the reproductive success of *Cardamine* and to possible explanations for low seed set.

*Cardamine* was usually found near first-order streams and seeps. The canopy was mostly closed with a vegetation composition representing that of a northern hardwood community or the transition zone between birch-spruce and spruce-fir. *Cardamine* was relatively rare in spruce-fir and mesic oak-beech forests. Associated plant species included mountain maple (*Acer spicatum*), red spruce (*Picea rubens*), yellow birch (*Betula alleghaniensis*), and buckeye (*Aesculus flava*). Although previously considered common in high-elevation boulderfields, we found *Cardamine* to be an infrequent component of these areas. The species was typically located in areas where direct herbaceous competition was low, microtopography was relatively well developed, and leaf litter did not accumulate. *Cardamine* usually occurred in local patches that were dispersed along a stream or seep, usually with locally dense moss, and where small openings in the canopy, shrub, and herbaceous layers permitted light to reach the forest floor for a portion of the day. This species was most frequently found rooted in moss or in moist rock crevices, though occasionally was found rooted in soil or streamside sandy depositions. Toward the western edge of its distribution, *Cardamine* was found on north-facing rock ledges and outcrops where seepage occurred and crevices in the rock were available for rooting. *Cardamine* appears to require moisture and does not respond well to shading by a dense herbaceous layer. Flowering was associated with surface water, low herb and shrub competition, and light gaps.

### *Rugelia nudicaulis*

*Rugelia nudicaulis* is a clonal, herbaceous perennial with a subterranean stem, or caudex. This species is a narrow endemic and is the only member of its genus. It has a basal whorl of leaves during the summer season that is replaced by a winter rosette near the end of October. This species flowers in late June through early August and fruits from July to September. When in





flower, *Rugelia* is visited by several species of bumblebees (*Bombus* spp.) and cuckoo bees (*Psithyrus* spp.), though it is not known which genera of insect contributes most significantly to its pollination. The fruits of *Rugelia* have numerous, hair-like pappus bristles that presumably aid in wind transport of the seed. Seed set and viability rates for *Rugelia* are unknown.

*Rugelia* was found to occupy narrow light gaps along ridgelines and gentle to moderately steep side-slopes. It was occasionally found along streams and seasonal drainages at mid-to-high elevations. *Rugelia* was most commonly found in both spruce-fir and northern-hardwood communities and in a broad range of habitat conditions ranging from dense regenerating fir to birch-maple-spruce with Pennsylvania sedge (*Carex pensylvanica*) present in the understory. The most frequently associated plant species were red spruce (*Picea rubens*), witch-hobble (*Viburnum lantanoides*), Fraser-fir (*Abies fraseri*), whorled wood aster (*Aster acuminatus*), southern lady fern (*Athyrium filix-femina* ssp. *asplenoides*), and mountain wood fern (*Dryopteris campyloptera*). *Rugelia* was found in a variety of topographic conditions, which partly explains the broad model. This species is often the dominant herb where it occurs. It has very large populations consisting of tens of thousands of crowns per hectare that are difficult to distinguish from other populations on the same ridge or mountain. In more exposed areas, *Rugelia* is often replaced by skunk goldenrod (*Solidago glomerata*), another endemic plant of the southern Appalachians. *Rugelia* was rarely found where there was dense shrub competition. Flowering was positively correlated with moderately sized light gaps and a canopy of mixed composition.

### *Glyceria nubigena*

*Glyceria nubigena* is a perennial grass that forms large clumps that can cover several square meters and be comprised of hundreds of stems. This species produces tall reproductive stems in mid-spring that flower in late May and June and produces seeds from mid-July to late August. *Glyceria* is probably wind-pollinated. There are no obvious adaptations for seed dispersal, although animal dispersal can not be ruled out given the presence of pronounced ridges on the lemma.



*Glyceria* was found to occupy clearings and rocky seeps in spruce-fir and northern hardwood communities. However, this species was very scarce in natural settings away from trails. This made assessment of its natural habitat difficult. It is likely that the range of *Glyceria* has expanded along trails, though it is not known whether this is a result of light availability, reduced competition, the availability of mineral soil, or a combination of factors. The majority of location documentation for *Glyceria* has been in artificial or semi-natural clearings at high elevations such as grassy balds, along trails, and near trailside shelters. Associated plant species were difficult to assign due to its trailside habitat, where numerous temporary species assemblages are found and delineation of habitat is difficult. Within its natural setting, best characterized as rocky seeps with an open canopy, *Glyceria* was most frequently associated with pink turtlehead (*Chelone lyonii*), Fraser fir (*Abies fraseri*), mountain cranberrybush (*Vaccinium erythrocarpum*), and red spruce (*Picea rubens*). Further efforts to define the natural habitat are needed.

#### *Calamagrostis cainii*

*Calamagrostis cainii* is a perennial, clump-forming grass. *Calamagrostis* flowers in June and July and produces fruit in July and August. A majority of the individuals in a population will flower in a given year. This species is probably wind-pollinated and wind-dispersed.

Locally, *Calamagrostis* is found in the unusual combination of rocky exposure with seepage. Regionally, it is found in exposed rocky seeps and rare high elevation rocky summits characterized by a conspicuous shrub component, low tree cover, and a relatively large amount of exposed rock (Wiser et al. 1997). It successfully roots in tiny fractures in the rock surface, forming dense mats of damp sod that apparently provide a rooting substrate for other rare herbaceous species of high elevation seeps. *Calamagrostis* is found in densely vegetated seeps only when overstory competition is low. The large populations are found in exposed areas where there is a sheet flow of water over the rock surface. The largest population of *Calamagrostis* dominates the herbaceous stratum. This species can also colonize moist, rocky areas along trails, especially near an abundant seed source (e.g., near the summit of Mount Le Conte, GSMNP). It



has been suggested that *Calamagrostis* is specific to the Anakeesta formation (slate, metasilt-stone), partly due to Anakeesta's landslide tendencies, which make open, rocky areas available for colonization. Although this is consistent with the majority of occurrences, the type location for *Calamagrostis* is found on more stable sandstone features in a steep, north-facing seep. This species frequently grows in steep or nearly vertical rocky areas and talus slopes with Carolina rhododendron (*Rhododendron carolinianum*), sand myrtle (*Leiophyllum buxifolium*), and deergrass (*Tricophorum cespitosum*). Other associated plant species include wretched sedge (*Carex misera*), linear-leaved gentian (*Gentiana linearis*), and mountain krigia (*Krigia montana*). Fertile stems of *Calamagrostis* were frequently associated with seepage and light availability.

### **Documented and potential threats**

#### Genetic diversity

Genetic diversity has been studied in *Rugelia*, *Glyceria*, and *Calamagrostis*. For *Rugelia*, allozyme diversity was very low compared with other endemic species, and *Glyceria* had no apparent allozyme diversity (Godt and Hamrick 1995). Their low genetic diversity may indicate a reduced capacity to adapt to changes in their environment.

Conversely, *Calamagrostis* showed levels of allozyme diversity, which were more comparable with rates for other narrow endemic plant species (Godt et al. 1996). However, additional analysis for *Calamagrostis* revealed a positive correlation between genetic diversity and population size (Godt et al. 1996). This finding could imply that the bulk of the genetic diversity is found in just three large populations, all of which are found on a single mountain summit in GSMNP (Mount Le Conte). If this is the case, one catastrophic event such as a large landslide could significantly reduce the genetic variation found within this species.

Additional genetic analyses need to be conducted for all four species to indicate the genetic similarity between populations, and to learn which populations represent the majority of the genetic variation.





### Air pollution

Ozone is an airborne pollutant with extremely elevated concentrations in the southern Appalachians. Whereas ozone concentrations typically fluctuate diurnally at lower elevations, the duration and overall intensity of high ozone episodes increases with elevation leading to more chronically-exposed vegetation at higher elevations. Fumigation experiments in GSMNP demonstrated that high ozone concentrations cause leaf tissue injury in at least 30 species of plants, including *Glyceria* and *Rugelia* (Neufield et al. 1992; Renfro 1989). Evidence of foliar injury also has been observed in the field for *Glyceria* and *Rugelia* (Renfro 1989). It is unknown to what degree elevated ozone may influence viability and population structure of any plant species.

GSMNP has the highest deposition of nitrogen and sulfur of all monitored sites in North America (Johnson and Lindberg 1992). Acid precipitation (from rain and snow) is five to 10 times more acidic (pH 4.5) than natural precipitation (5.0-5.6) (source: NADP 1998). Fog is even more concentrated, with pH as low as 2 (source: EPA 1998). Effects of acid deposition and input of additional nitrogen on the high elevation ecosystems are adverse and well documented, but there is no current documentation of how these factors affect the study species.

### Introduced species

The introduced European wild boar (*Sus scrofa*) is known to cause serious damage to the herbaceous layer in high elevation forest communities (Bratton 1974, 1975). Boar rooting also has been shown to accelerate leaching of soil nutrients and to alter nitrogen transformation processes (Singer et al. 1984). Direct and indirect effects of boar rooting on distribution, abundance, and population dynamics of the study species have not been directly assessed, though we have observed small-scale destruction of *Rugelia* populations by boar activity.

The red spruce-Fraser fir forest ecosystem is in serious decline due to the infestation of Fraser fir by the introduced balsam woolly adelgid (*Adelges piceae*) (White 1984). First detected in the southern Appalachians in 1957 (Speers 1958), it has since impacted all known stands of Fraser fir.



A loss of genetic diversity and eventual extinction is feared. Widespread loss of the fir canopy has altered the pattern of succession at higher elevations by allowing more sunlight to penetrate to the forest floor, thereby increasing temperature and reducing relative humidity in the understory (DeSelm and Boner 1984). These changes in microclimate have tended to favor the success of shrub species, such as blackberry (*Rubus canadensis*), over that of herbaceous species (DeSelm and Boner 1984, Nicholas et al. 1992). Although *Rugelia* has been observed to respond positively to small-scale clearings and light gaps, it is currently unknown how the loss of Fraser fir will affect landscape-wide distribution and abundance of this narrow endemic. Monitoring data suggest that *Rugelia* can persist in a vegetative condition for long periods in deep shade following the death of the fir canopy and subsequent growth of dense stands of regenerating fir. We do not have any information on the response of the other three study species to Fraser fir death.

#### NPS activities

Road and trail maintenance also presents a potential threat to populations of *Rugelia*, *Glyceria*, and *Calamagrostis* because a large proportion of the total number of occurrences for these species is found along trails or roadways. Brush clearing and mowing are conducted in GSMNP in the interest of providing safe roadway shoulders and enabling ease of movement along trails. Unfortunately, these activities are typically performed during the summer months when these species are particularly vulnerable because most of their stored energy has been redirected toward reproductive stems. *Rugelia* has been observed to have a higher rate of flowering and seed production along trails – presumably in response to higher light levels – so trail maintenance activities may unequally affect reproductive individuals of this species, albeit “positively”. The tall reproductive stems of *Glyceria* (up to 2 meters or 7 feet) are frequently severed completely during the course of summer trail and road maintenance. This eliminates the possibility of reproduction for the affected individuals for at least that season. As much as 25% of all fertile clumps are cut back prior to seed maturation each year (K. Langdon, unpublished data, GSMNP). To what degree successive cuttings by trail maintenance decrease survivorship of individual

The first part of the year was spent in the  
study of the history of the country and the  
character of the people. The second part was  
spent in the study of the natural history of the  
country and the character of the people. The third  
part was spent in the study of the political  
history of the country and the character of the  
people. The fourth part was spent in the study  
of the social history of the country and the  
character of the people. The fifth part was  
spent in the study of the economic history of  
the country and the character of the people.

The first part of the year was spent in the  
study of the history of the country and the  
character of the people. The second part was  
spent in the study of the natural history of the  
country and the character of the people. The third  
part was spent in the study of the political  
history of the country and the character of the  
people. The fourth part was spent in the study  
of the social history of the country and the  
character of the people. The fifth part was  
spent in the study of the economic history of  
the country and the character of the people.

clumps of *Glyceria* is not known. Since over 90% of currently known occurrences of *Glyceria* occur along trails and roads, this topic requires further study.

### Data gaps

We have little data on population trends for the four study species. One population of *Cardamine* has been monitored since 1992 without any significant changes observed. Two occurrences of *Rugelia* also are monitored. The plots for these two occurrences were placed in areas formerly dominated by Fraser fir. In just the last eight years, these plots have undergone significant changes in the structure of the understory since the death of the fir canopy (circa 1992). No significant changes in rosette number have been detected during 10 years of monitoring, though a gradual decrease in plant size, flowering, and vigor has been noted. *Glyceria* and *Calamagrostis* are not currently monitored, and additional monitoring of all four species will be necessary to resolve their responses to changes in high-elevation forest communities.

Changes in the composition and structure of the vegetative communities that support the study species may be the most direct factors that will affect distribution and abundance in the future, though these types of changes are most difficult to predict due to complexity of the systems involved. Global climate change is another factor that may shift species niche space and relative abundance in this century. Some climatic change models predict an average global temperature increase of 5° over 100 years, with significantly altered temperature regimes and patterns of rainfall worldwide.

Present-day abundance and distributions of *Cardamine*, *Rugelia*, *Glyceria*, and *Calamagrostis* likely were influenced by human activity in the past century. Intensive logging of high-elevation communities through the 1930s, reduced the extent of spruce-fir forests in the southern Appalachians by 50% (Pyle 1984) and was known to cause extensive fires and severe erosion (Korstian 1937). At this point we only can speculate how the combination of habitat loss, a





shifting landscape of suitable habitat, non-native infestations, atmospheric pollutant deposition, and forest succession will interact to influence the long-term viability of the study species.

### Model output and evaluation

Summary statistics for the four study species are shown in Table 3. Since  $D^2$  is a unitless measure of deviation from an optimal combination of variables, “small values” represent habitat conditions most favorable for the species. All summaries and analyses of the model output are specific to the study area within GSMNP, defined as all Park property above 914 meters (3,000 feet) elevation.  $D^2$  output for each species ranged from under five to over 1,000. Average values for the study area ranged from 36.18 for *Rugelia* to 659.86 for *Calamagrostis*. Small average values mean that suitable habitat is relatively common in the study area; large values mean suitable habitat was relatively uncommon. Given the results, *Rugelia* had the broadest extent of suitable habitat, whereas suitable habitat for *Calamagrostis* was uncommon.

**Table 3.** Statistical summary for Mahalanobis distance ( $D^2$ ) output.

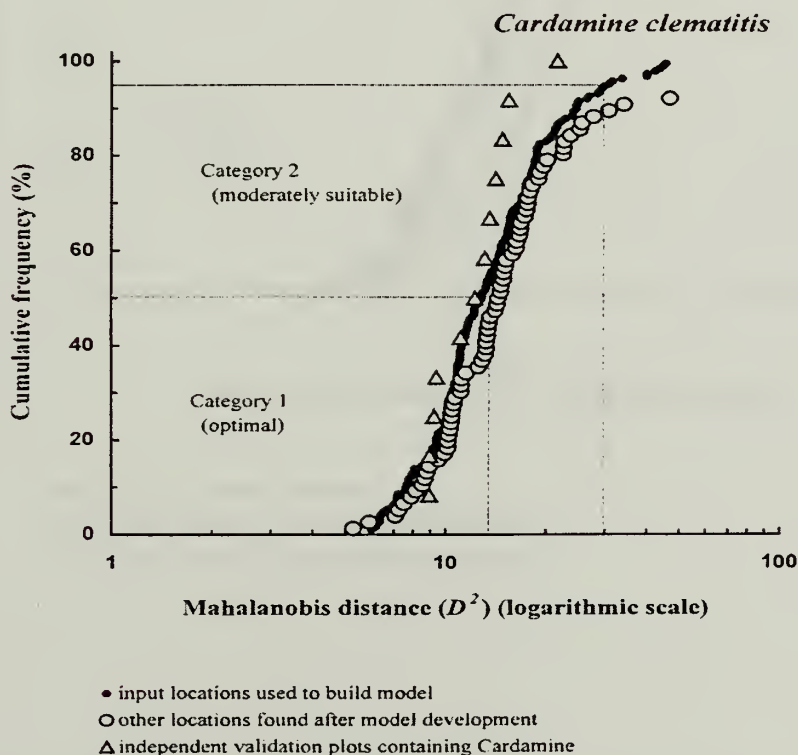
Species	$D^2$ min	$D^2$ max	mean	SD
<i>Cardamine</i>	3.75	1,164.90	53.95	38.78
<i>Rugelia</i>	1.78	2,134.39	36.18	19.13
<i>Glyceria</i>	1.69	3,106.06	118.75	77.78
<i>Calamagrostis</i>	3.33	4,056.81	659.86	394.97

Cumulative frequency of  $D^2$  values corresponding to known occurrences for each species is depicted in Figures 1 through 4. Distribution of  $D^2$  values is graphed on a logarithmic scale. A vertically-oriented distribution curve indicates a tighter distribution of values, whereas horizontal tendencies indicate a broader distribution. Habitat categories for each model were defined using the range of values for positions used to build the model (see Methods).  $D^2$  values defining the upper limits for category 1 and 2, respectively, were as follows: *Cardamine*, 12.8 and 31 (Figure 1); *Rugelia*, 11.1 and 27 (Figure 2); *Glyceria*, 10.3 and 28.5 (Figure 3); and *Calamagrostis*, 20



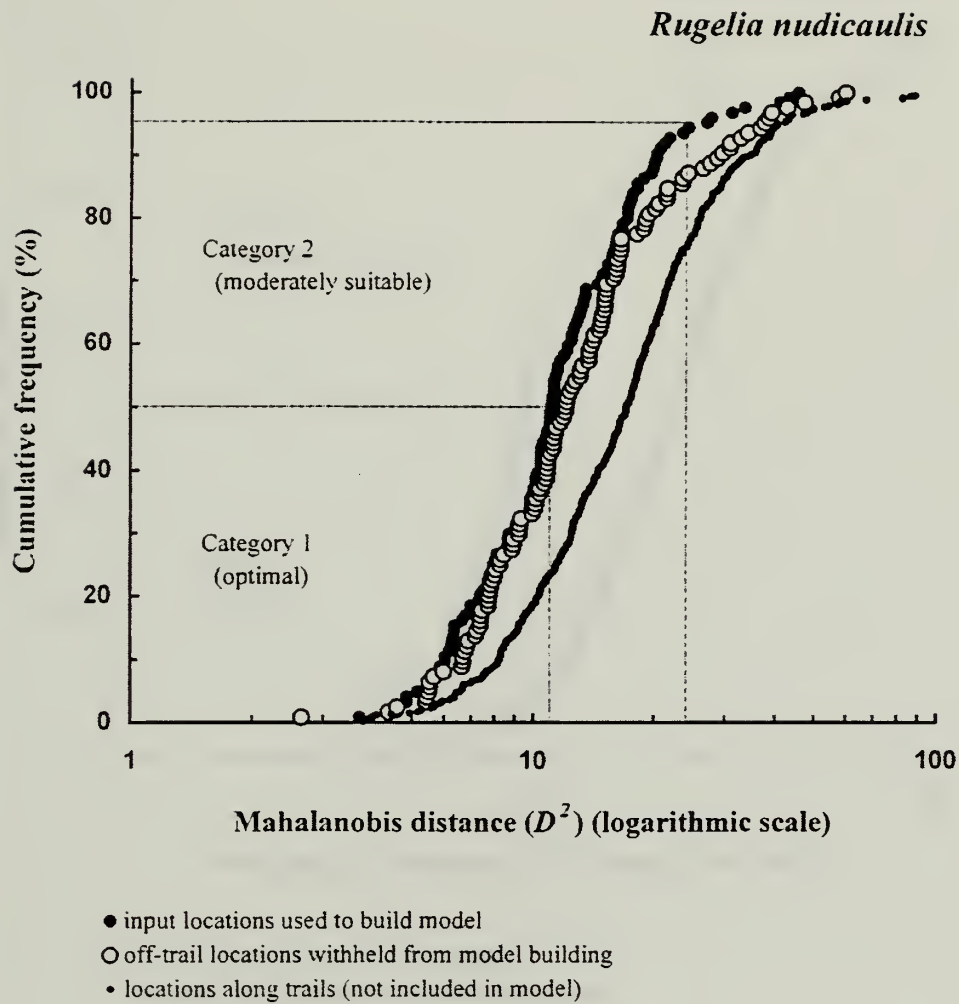


and 42 (Figure 4). Vertical lines in Figures 1 through 4 indicate the  $D^2$  cutoffs defining each habitat category, whereas horizontal lines indicate the percentage of known occurrences included within that range of values. These graphs are intended to convey the distribution of  $D^2$  values corresponding to positions that went into creating the models ("input locations" in Figures 1 through 4). The distribution curves for *Cardamine* (Figure 1), *Rugelia* (Figure 2) and *Glyceria* (Figure 3) all show a similar sigmoidal shape, indicating a normal distribution with a frequency maximum toward the median value (i.e., that associated with the 50<sup>th</sup> percentile on the y-axis in these graphs).



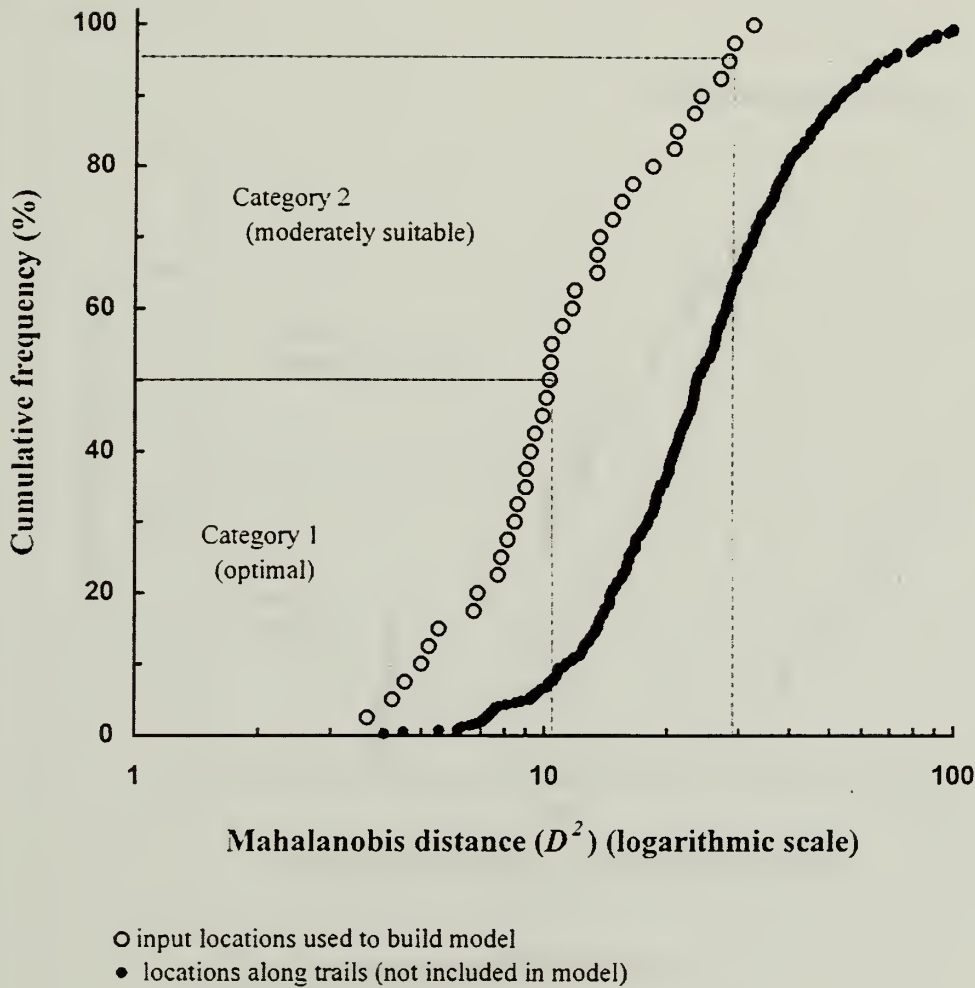
**Figure 1.** Cumulative frequency for  $D^2$  values for *Cardamine clematitis* positions.





**Figure 2.** Cumulative frequency of  $D^2$  values for *Rugelia nudicaulis* positions.



*Glyceria nubigena*

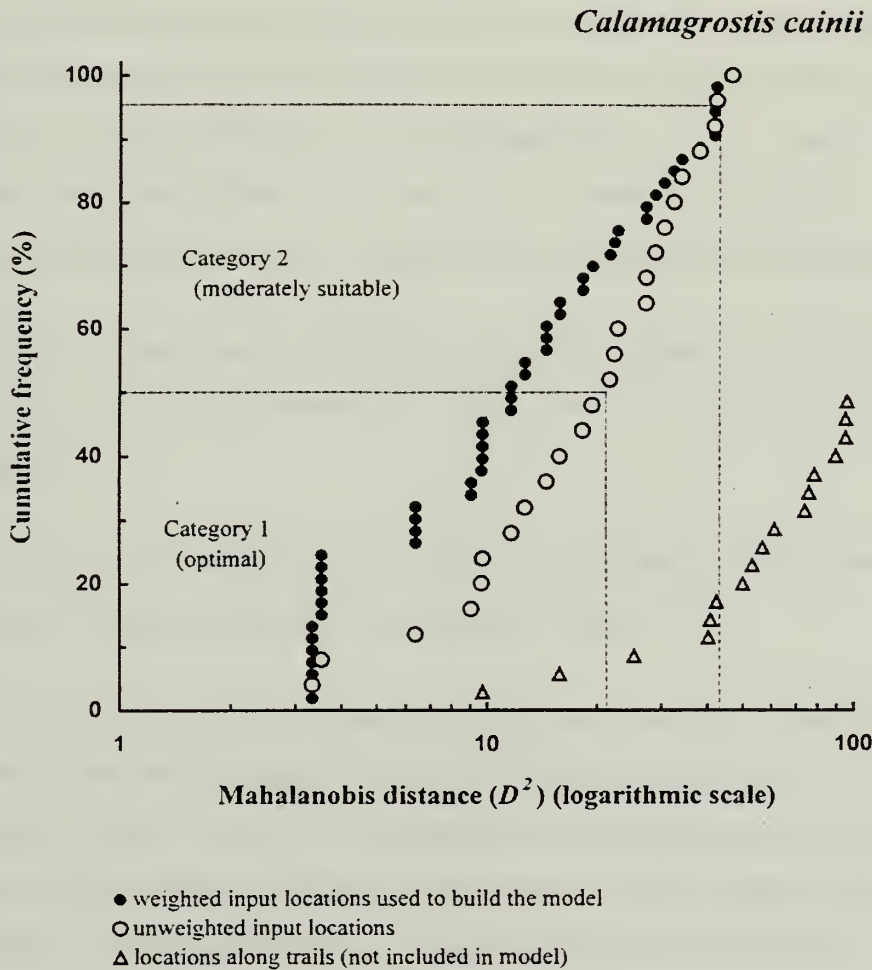
**Figure 3.** Cumulative frequency of  $D^2$  values for *Glyceria nubigena* positions.

In contrast, *Calamagrostis* had irregular distribution curves (Figure 4). Both weighted and unweighted input sets for *Calamagrostis* are depicted. The difference between these two curves reflects the influence of the weighting procedure on the model. The weighted data, upon which the model was based, had a greater frequency at low values (Figure 4) which indicates that positions with the most weight (i.e., those where *Calamagrostis* was most abundant) also have





relatively low  $D^2$  values associated with them. We selected the unweighted 50% cutoff because it corresponded to a higher value of  $D^2$ ; consequently, more area was included in the “optimal” habitat category.



**Figure 4.** Cumulative frequency of  $D^2$  values for *Calamagrostis cainii* positions.

Model values corresponding to positions that were either withheld during model development (as for *Rugelia*) or encountered after model development (as for *Cardamine*) corresponded closely with values for model inputs. Cumulative frequency curves for these positions (“other locations” and “off-trail locations” Figures 1 and 2, respectively) show a similar shape and position along the ordinate axis as the model inputs.



Each curve fits the habitat categories well, with approximately 40% of positions fitting into Category 1 habitat and over 90% of the positions fitting into categories 2 and 3. This indicates that the models for *Cardamine* and *Rugelia* performed well in predicting the occurrence of these species. Both models were based on a relatively large number of independent coordinates (187 and 124, respectively) which appears sufficient to delineate suitable habitat. In contrast, the models for *Glyceria* and *Calamagrostis* were based on a limited number of high-quality, off-trail location coordinates (*Glyceria*,  $n = 40$ ; *Calamagrostis*,  $n = 25$ ). As such, we were unable to withhold any subset of these model positions for independent evaluation. We currently have few criteria for evaluation of these two models other than their fit with original data.

Trailside occurrences were not predicted as well as positions away from trails. The best fit for trailside occurrences was observed for *Rugelia*, for which over 80% fit into categories 2 and 3 (Figure 6).

For *Glyceria*, only 10% of trailside occurrences fit category 3, and less than 70% of the distribution fit into categories 2 and 3 combined (Figure 7).

The fit was even worse for *Calamagrostis*, with lower than 20% of trailside occurrences fitting into the highest habitat categories (Figure 8). Recall that categories 2 and 3 were based on the  $D^2$  values below which 50% and 95% (respectively) of off-trail input locations occurred. A low level of fit with the input data may indicate that a species occupies different habitat conditions along trails than it does away from trails. There are several potential explanations for these results: (1) small, trailside occurrences for *Calamagrostis* may be associated with small, rocky areas missed during photographic interpretation or below the resolution of the data; (2) a mass effect may exist whereby large neighboring populations producing an overabundance of seed allow for the dispersal of *Calamagrostis* into marginal habitat; (3) trailside habitat may represent a set of continually-disturbed conditions that allow micro-disturbance species such as *Rugelia*, *Glyceria* and *Calamagrostis* to repeatedly establish themselves in a broader range (though less natural) of habitat conditions; and (4) the models for *Calamagrostis* and *Glyceria* may be based on too few



positions to adequately describe the full range of habitat conditions for these species.

#### Frequency of suitable habitat

The frequency of suitable habitat in GSMNP is summarized for each species in Table 4. When totaled, categories 1 and 2 comprised over 44,000 hectares for *Cardamine*, 41,000 hectares for *Rugelia*, 10,000 hectares for *Glyceria*, and 300 hectares for *Calamagrostis*. These numbers indicate how much “potential” habitat exists in GSMNP, but they do not indicate how frequently each species occurs within that potential habitat. Likely, these species are much more rare than their potential habitat. This rarity might be attributed to competition between plant species or limited dispersal between patches of suitable habitat. In order to estimate habitat saturation for each species, it is necessary to visit numerous locations within the predicted areas.

**Table 4.** Frequency of suitable habitat in GSMNP.

Species	Habitat category <sup>a</sup>	Number of pixels	Habitat frequency (hectares)	Percent of study area <sup>b</sup>
<i>Cardamine</i>	1	54,360	4,892.4	4.03
	2	437,353	9,361.8	32.39
	total (1 and 2)	491,713	14,254.2	36.42
<i>Rugelia</i>	1	61,906	5,571.5	4.58
	2	400,167	36,015.0	29.63
	total (1 and 2)	462,073	41,586.5	34.21
<i>Glyceria</i>	1	11,889	1,070.0	0.88
	2	105,973	9,537.6	7.85
	total (1 and 2)	117,862	10,607.6	8.73
<i>Calamagrostis</i>	1	150	13.5	0.01
	2	3,274	294.7	0.24
	total (1 and 2)	3,424	308.2	0.25

<sup>a</sup> Category 1 or “optimal” habitat was defined to capture half the range of  $D^2$  corresponding to input positions. Category 2 (“moderately suitable habitat”) captured the next 45% of that range.





<sup>b</sup> The study area was defined as GSMNP above 914 meters (3,000 ft) elevation. The study area was 121,081 ha (1,350,345 pixels at 0.09 ha/pixel).

The geographic distribution of suitable habitat for the study species is depicted in Maps 5 through 14. Each model depicts habitat as delineated by the habitat categories. Close-ups of either Clingmans Dome or Mount Le Conte are provided so the output can be seen at a larger scale. In each case, positions used to build the models are superimposed to show the fit with known occurrences.

The *Cardamine* model identified high elevation drainages in central and eastern GSMNP as optimal habitat (Maps 5 and 6). Suitable habitat was patchy and isolated in distribution, being separated by high elevation ridgelines (which were identified as unsuitable habitat).

Suitable habitat for *Rugelia* was much more contiguous and concentrated (Maps 7 and 8), with optimal habitat largely confined to approximately ten large areas associated with high elevation summits. Optimal habitat for this species was most concentrated in central GSMNP in the vicinity of Clingmans Dome.

Suitable habitat for *Glyceria* (Maps 9 and 10) was much less frequent overall, with optimal habitat occurring in small, patchy clusters that have a southerly exposure and high slope position.

Finally, *Calamagrostis* habitat was highly restricted to a few, very small areas near the summit of Mount Le Conte (Maps 11 and 12). Optimal habitat had a southerly exposure and high slope position.

#### *Cardamine clematitis* field validation

The presence/absence results for the *Cardamine* validation plots are summarized in Table 5. *Cardamine* was encountered in 12 out of 120 (10%) tenth-hectare plots stratified across the full range of  $D^2$  values. These 12 occurrences, represented by green dots in Map 1, matched well with other known occurrences. The shape of the curve and its distribution along the ordinate axis





is similar to that for the input positions as well as for the set of positions encountered after model development (Figure 1). Logistic regression of validation data, with presence/absence as the dependent variable and  $D^2$  as the independent variable, showed a significant inverse relationship (parameter estimate = -0.1323,  $P = 0.0394$ ).

**Table 5.** Presence/absence of *Cardamine clematitis* in validation plots, GSMNP.

Stratification Class	Min $D^2$	Max $D^2$	Area (ha)	<i>Cardamine</i> present	<i>Cardamine</i> absent
1	3.748	10.514	1,921	4	16
2	10.514	13.361	3,843	3	17
3	13.361	17.291	7,686	4	16
4	17.291	23.981	15,372	1	19
5	23.981	39.676	30,745	0	20
6	39.676	1164.900	61,490	0	20
<b>Total</b>			<b>121,357</b>	<b>12</b>	<b>108</b>

Chi-square analysis of presence in stratification classes supported the significance of the inverse relationship (Table 6). These results confirm that the model is useful for depicting suitable habitat for *Cardamine*.

**Table 6.** Chi-square analysis of *Cardamine* presence in validation plots, as grouped by validation classes.

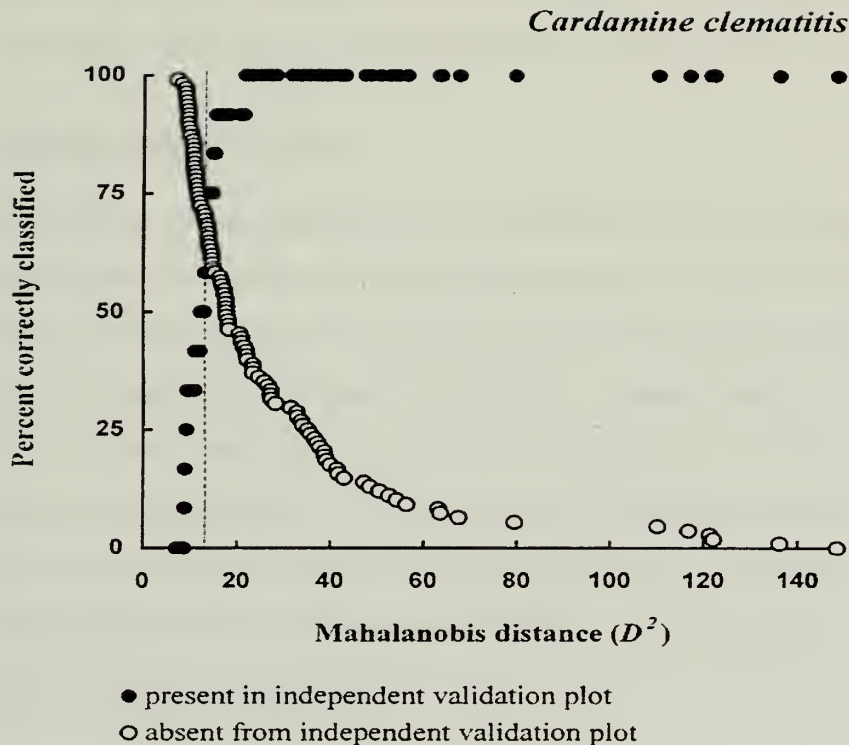
Validation Class Groupings <sup>a</sup>	Effective range for $D^2$ by grouped classes	Chi-square	df	P-value
1, 2, 3, 4, 5, 6	same as stratification classes in Table 5.	9.0	5	< 0.10
1 - 2, 3 - 4, 5 - 6	3.748 - 13.361; 13.361 - 23.981; 23.981 - max	6.5	2	< 0.05
1 - 3, 4 - 6	3.748 - 17.291; 17.291 - max	8.3	1	< 0.005

<sup>a</sup> Analysis was on grouped classes (Table 5), which were combined for analysis as delineated by commas.



We attempted to determine a suitable cutoff for  $D^2$  that maximized the model's ability to predict both presence and absence of *Cardamine*.  $D^2$  values associated with validation plots provided a means of evaluating a range of possible cutoffs. Correct classification was defined as presence of *Cardamine* in plots with  $D^2$  below a given cutoff, and absence of *Cardamine* in plots with  $D^2$  above the cutoff. When the percentage of plots with correct classifications is charted against the range of possible  $D^2$  cutoffs, the curve for correctly-classified presence increases with  $D^2$ , and the curve for correctly-classified absence decreases with increasing  $D^2$  (Figure 5). This is consistent with overall correlation of *Cardamine* occurrence and low values of  $D^2$ . As the  $D^2$  cutoff increases, more of these occurrences are correctly classified until a maximum of 100% is reached (above 22.3 in Figure 5). In contrast, fewer of the plots in which *Cardamine* was absent are correctly classified as the  $D^2$  cutoff is increased. The point at which the two curves intersect represents the best cutoff, that is, where correct classification of both presence and absence is maximized. The intersection occurs at a  $D^2$  value of 13.81, at which point correct classifications of both presence and absence is 66.7%. Important to note, is that when defining suitable habitat by this cutoff, one third of the occurrences in the validation plots are, in effect, misclassified. While a more inclusive cutoff at a larger  $D^2$  value would be more sensitive to *Cardamine* presence, it would be less specific because it would also include a much larger proportion of habitat in which the species is absent. Therefore, the 13.81 cutoff represents the best definition of suitable habitat for *Cardamine* given available data.





**Figure 5.** Classification of validation data for *Cardamine clematitis*.

#### *Cardamine clematitis* suitable habitat

The distribution of suitable habitat for *Cardamine* is depicted in Maps 13 and 14. These maps represent habitat as defined by field validation rather than range of values for input locations (as in Maps 5 and 6). Areas with  $D^2$  below 13.81 comprise approximately 6,556.6 hectares, or 5.4% of the study area. However, *Cardamine* was found in only 18% of validation plots below this cutoff. This may be a result of a combination of factors including dispersal limitation between habitat patches, competition with other species, and stochastic local extinction events. If this habitat-saturation rate is representative, the extent of area in which *Cardamine* is likely to be found is approximately 1180.1 hectares, or just under 1% of the study area. Whenever *Cardamine* was found, it only covered a very small area, never exceeding even 1% of a tenth-hectare quadrat (10 square meters). So although *Cardamine* is predicted to occur in approximately 12,000 tenth-hectare parcels across GSMNP, total cover is probably only a tiny





fraction of that area – perhaps on the order of one to 10 hectares total. This is the best information presently available on the distribution and relative frequency of *Cardamine* in GSMNP.

#### *Cardamine clematitis* summary

Field validation of the *Cardamine* model enabled the following analyses and evaluations: (1) we confirmed that *Cardamine* occurrence corresponds to low values of  $D^2$ , and that the model was useful in delineating habitat for this species; (2) we were able to establish a functional cutoff for  $D^2$ , which makes the model much more useable to resource managers; and (3) we now understand that *Cardamine* is about five times as rare as its habitat, and may be expected to be present at some level in approximately 1,200 hectares across GSMNP (though it is likely to cover much less than 1% of that area). We emphasize field validation is a critical component of model evaluation and should be performed for the *Rugelia*, *Glyceria*, and *Calamagrostis* models before they are applied to resource management.

#### **Suggestions for model improvement**

Our habitat suitability models were built using remotely-sensed GIS data. This is advantageous because the models can be applied to large geographic areas, making them more applicable to conservation and management of target species. However, due to the limitations of remotely-sensed data, our models do not incorporate data on geology, soil, or vegetation structure and composition. This environmental data would probably have a large influence on species occurrence. For example, the availability of data on soil type is extremely limited at this time and the current Park vegetation map was developed using satellite imagery to “coarsely” classify plant communities (i.e., nine forest types and three herbaceous/shrub communities) (MacKenzie 1993). Currently, over 70 different plant communities are estimated to occur in the Park. With partners such as USGS and TNC, efforts are underway to map both soil type and plant communities Park-wide. This will enable the development of more powerful and accurate habitat models in the near future.



Our model for *Cardamine* successfully predicted the occurrence of new populations in GSMNP and, therefore, has valuable conservation applications. However, use of this model has some constraints in that *Cardamine* was also frequently absent from areas identified as suitable habitat. At present we cannot say whether the absence from suitable areas is because we neglected to include an important landscape-scale variable during model construction, or because the 30-meter GIS variables only go so far in specifying the occurrence of rare understory herbs. Fine-scale information on the structure and composition of the herbaceous vegetation might further explain absence from areas identified as suitable habitat. There also may be a dispersal component to habitat occupancy, whereby suitable areas distant from extant populations have a reduced likelihood of successful seed dispersal and establishment. Further work toward improving the model for *Cardamine* should be directed toward investigating these relationships.

Model improvement for *Calamagrostis* and *Glyceria* should focus primarily on increasing the number of input coordinates. Because each of these models was based on a small number of input coordinates, the range of habitat conditions represented by the model input may have been too narrow for predicting Park-wide distribution. Development of satisfactory models for these species may require an overall improvement in resolution of digital data, especially of exposed, rocky areas for *Calamagrostis*, and of canopy structure for *Glyceria*.

Ultimately, habitat models might incorporate information on the demographics of the target species, such as growth rates, reproduction, and mortality. The addition of explicit components for dispersal and stochastic disturbance would increase our ability to predict changes in population structure. However, these enhancements would make the models more complex and would require a large amount of data to estimate the demographic parameters. Nevertheless, such enhancements may be necessary to estimate the demographic patterns of these species in response to environmental changes and to determine whether the target species are stable in the environment.



## CONCLUSION AND RECOMMENDATIONS

Our field surveys identified suitable habitat and discovered new occurrences for each of the four study species. This is the first time that the Mahalanobis distance statistic ( $D^2$ ) has been used to model rare plant distribution. Our efforts produced a valid model depicting the distribution of *Cardamine* habitat in GSMNP (Maps 13 and 14). The models for *Cardamine* and *Rugelia* were found to adequately predict other post-modeling occurrences. However, the models for *Rugelia*, *Glyceria*, and *Calamagrostis* require field validation for better evaluation of their predictive power.

### *Cardamine clematitis*

*Cardamine* has the broadest geographic range, and most well-defined habitat, of the four study species. Our modeling and validation efforts show that habitat for *Cardamine* is patchy and dispersed throughout the high elevations of central and eastern GSMNP. Habitat occupancy rates are low which suggest that further efforts to refine the model using fine-scale habitat data are needed. Approximately 45 documented populations of *Cardamine* occur in 18 counties in North Carolina, Tennessee and Virginia; at least 18 populations occur inside GSMNP and at least 27 outside the Park. It is likely that additional populations will be located with further surveys. We recommend the current rank of G2G3.

### *Rugelia nudicaulis*

*Rugelia* is only found within GSMNP. It is documented from two counties in North Carolina and two counties in Tennessee. Population boundaries are extremely difficult to assign for *Rugelia* given its expansive, continuous occurrence on mountain summits and side-slopes. The GIS habitat model indicates that suitable habitat is largely confined to the higher ridges and mountain summits in central and eastern GSMNP. The known distribution of *Rugelia* closely resembles the distribution of suitable habitat as predicted by the model. However, this model has not been tested and field validation is required for better interpretation of its predictive capabilities.





Additional work on the genetic differences between populations and on the response of *Rugelia* to changes in overstory composition is needed before a change in species status can be recommended. We recommend that state rank and status be re-evaluated and that the species be assigned a global rank of G2.

### ***Glyceria nubigena***

Presently *Glyceria* is documented from two counties in Tennessee and three counties in North Carolina. The majority of occurrences are found in artificial clearings along trails and roadways. Our habitat model for this species was developed using only 40 off-trail positions, which may be too small a number to produce an adequate model. Additional off-trail occurrences should be sought and used to produce a more robust model, which should be subsequently field tested. Preliminary evidence suggests that the genetic diversity of *Glyceria* is very low, which might make it more vulnerable to environmental fluctuations. Estimation of population boundaries is difficult for this species and its natural habitat requirements are not well understood. Additional work is needed to document *Glyceria*'s natural habitat requirements and its response to changes in forest structure and composition. Further work is needed to refine criteria for distinguishing populations prior to any change in species status. We recommend that trail maintenance at the high elevations in GSMNP be timed to avoid damaging the reproductive stems during the late summer months. By not clearing vegetation from high elevation trails during this period (i.e., late May through mid-September), individual trailside populations would be allowed to flower and produce seed successfully. We recommend the current rank of G2.

### ***Calamagrostis cainii***

*Calamagrostis* is known from one county in Tennessee and two counties in North Carolina. It is restricted to just three mountain summits and is the rarest of the four species investigated. Our habitat model indicates that *Calamagrostis* habitat is extremely rare in GSMNP and that such habitat is highly clustered around Mount LeConte. However, this model was based on a small number of independent coordinates and requires field validation before it can be fully evaluated.





Only five populations of *Calamagrostis* have been documented in the world, three of which are located in GSMNP within a few miles of Mount Le Conte's summit. A majority of the total number of individuals of this species is found in only one population. And, there is a preliminary documentation of below-normal genetic diversity and a purported correlation between low diversity and small population size. This may indicate a reduced capacity for *Calamagrostis* to respond to environmental fluctuations, and, that the bulk of species-wide genetic diversity may be found in just one or two populations. We recommend that *Calamagrostis* be considered for federal listing as Threatened or Endangered given its extreme rarity and additional searches for populations outside GSMNP should be conducted. Its global rank of G1 should be retained.

## SUMMARY

This is the first time 30-meter resolution data have been applied to model distribution of rare plant species. The development of predictive distribution models is an important component for long-term conservation planning. By understanding the distribution of suitable habitat, systematic searches can be made more efficient and potential habitat can be identified for either restoration or re-introduction of extirpated species. The distribution and connectivity of potential habitat also has serious implications for dispersal probability and, consequently, for the long-term survival of a species given changes in ecosystem function. Such habitat characteristics can only be evaluated with the aid of a habitat-suitability model. However, the theoretical basis for interpretation of these models is still not well developed or understood. For example, we do not know how to properly interpret the low-habitat occupancy rates observed for *Cardamine*, despite the fact that the *Cardamine* model was otherwise good at depicting habitat. Clearly, additional research into theoretical and applied aspects of habitat delineation and niche occupancy is needed to properly interpret application of habitat models to rare plant conservation.



## LITERATURE CITED

- Al-Shehbaz, I. A. 1988. The genera of Arabideae (Cruciferae; Brassicaceae) in the southeastern United States. *Journal of the Arnold Arboretum*, 69: 85-166.
- Amoroso, J. L., ed. 1997. Natural heritage program list of the rare plant species of North Carolina. North Carolina Natural Heritage Program, N.C. Dept of Environment, Health and Natural Resources, Raleigh. 88 p.
- Beers, T. W., P.E. Dress, and L. C. Wensel. 1966. Aspect transformation in site productivity research. *Journal of Forestry* 64: 691-692.
- Beven, K. J., and M. J. Kirby. 1979. A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.* 24: 43-69.
- Bratton, S. P. 1974. The effect of the European wild boar (*Sus scrofa*) on high-elevation vernal flora in the Great Smoky Mountains. *Bull. Torrey Bot. Club* 101: 198-206.
- Bratton, S. P. 1975. The effect of the European wild boar (*Sus scrofa*) on gray beech forest in the Great Smoky Mountains. *Ecology* 56: 1356-1366.
- DeSelm, H. R. and R. R. Boner. 1984. Understory changes in spruce-fir during the first 16-20 years following the death of fir. P. 51-69. *In*: White, P.S. (ed.). The southern Appalachian spruce-fir ecosystem: its biology and threats. U. S. Department of the Interior, National Park Service, Research/Resources Management Report SER-71. 268 p.
- Dudley, T. R. 1974. The correct authority for *Cardamine clematitis* (Cruciferae). *Rhodora* 76: 53-57.
- Duncan, W. H. and J. T. Kartesz. 1981. Vascular flora of Georgia: An annotated checklist. University of Georgia Press, Athens.
- Environmental Protection Agency. 1997. Clean Air Status and Trends Network (CASTNet). Mountain Acid Deposition Program: Data report – 1996 monitoring season. QST Environmental, Inc. 52 p.
- Fernald, M. L. 1950. Gray's manual of botany, 8th ed. Dioscorides Press, Portland, Oregon. 1632 p.
- Gaddy, L. L., C. C. Douglass, A. E. Hodge, H. LeGrand. 1984. Plants new to South and North Carolina. *Castanea* 49 (2): 97.



- Gleason, H. A. 1952. The new Britton and Brown illustrated flora of the northeastern United States and adjacent Canada, 3 vols. Hafner Press, New York.
- Gleason, H. A. and A. Cronquist. 1991. Manual of vascular plants of northeastern United States and adjacent Canada, 2nd ed. New York Botanical Garden, Bronx, NY 10458. 910 p.
- Godfrey, R. K. and J. W. Wooten. 1981. Aquatic and wetland plants of southeastern United States: Dicotyledons. University of Georgia Press, Athens. 933 p.
- Godt, M.J. W. and J. L. Hamrick. 1995. Allozyme variation in two Great Smoky Mountain endemics: *Glyceria nubigena* and *Rugelia nudicaulis*. Journal of Heredity 86: 194-198.
- Godt, M.J. W., Johnson, B. R. and J. L. Hamrick. 1996. Genetic diversity and population size in four rare Southern Appalachian plant species. Conservation Biology 10 (3): 796-805.
- Halpin, P. N. 1995. A cross-scale analysis of environmental gradients and forest patterns in the Giant Sequoia - Mixed Conifer forest of the Sierra Nevada. Ph.D. dissertation, University of Virginia, Charlottesville.
- Hosmer, D. W., Jr., and S. Lemeshow. 1989. Applied logistic regression. John Wiley and Sons, New York, N. Y. 307 p.
- Johnson, D. W. and S. E. Lindberg. 1992. Atmospheric Deposition and Forest Nutrient Cycling: a Synthesis of the Integrated Forest Study. Springer-Verlag, New York.
- Kartesz, J. T. 1994. A synonymized checklist of the vascular flora of the United States, Canada, and Greenland, 2nd ed. Timber Press. Portland, OR. 573 p.
- Korstian, C. F. 1937. Perpetuation of spruce on cut-over and burned lands in the higher southern Appalachian Mountains. Ecological Monographs 7: 125-167.
- Ludwig, J. C. 1997. Natural Heritage Resources of Virginia: rare vascular plants. Natural Heritage Technical Report 97-1. Virginia Department of Conservation and Recreation, Division of Natural Heritage, Richmond, Virginia. 34 p. plus appendices.
- MacKenzie, M. D. 1993. The vegetation of Great Smoky Mountains National Park: past, present and future. Ph.D. Dissertation, University of Tennessee, Knoxville, TN.
- Marcot, B. G., M. G. Raphael, and K. H. Berry. 1983. Monitoring wildlife habitat and validation of wildlife-habitat relationships models. Trans. North Am. Wildl. and Nat. Resour. Conf. 48: 315-329.





- McNab, W. H. 1989. Terrain shape index: quantifying effect of minor landforms on tree height. *Forest Sci.* 35: 91-104.
- McNab, W. H. 1993. A topographic index to quantify the effect of mesoscale landform on site productivity. *Canadian Journal of Forest Research* 23: 1100-1107.
- Miller, R. 1986. Predicting rare plant distribution patterns in the southern Appalachians of the southeastern U.S.A. *Journal of Biogeography* 13: 293-311.
- NADP. 1998. National Atmospheric Deposition Program. Illinois State Water Survey. Champaign, IL. 15 p.
- Neufield, H. S., Renfro, J. R., Hacker, W. D. and D. Silsbee. 1992. Ozone in Great Smoky Mountains National Park: Dynamics and effects on plants. In *Tropospheric Ozone and the Environment II: Effects, modeling and control*. R.L. Berglund, editor. Air and Waste Management Association, Pittsburg, PA.
- Nicholas, N. S., S. M. Zedaker, and C. Eagar. 1992. A comparison of overstory community structure in three southern Appalachian spruce-fir forests. *Bull. Torrey Bot. Club* 119: 316-332.
- Nordman, C. 1996. Tennessee Natural Heritage Program: rare vascular plant list. Tennessee Department of Environment and Conservation, Division of Natural Heritage, Nashville.
- Parker, A. J. 1982. The topographic relative moisture index: an approach to soil-moisture assessment in mountain terrain. *Phys. Geogr.* 3(2): 160-168.
- Pyle, C. 1984. Pre-park disturbance in the spruce-fir forests of Great Smoky Mountains National Park. Pp. 115-130. In: White, P.S. (ed.). *The southern Appalachian spruce-fir ecosystem: its biology and threats*. U. S. Department of the Interior, National Park Service, Research/Resources Management Report SER-71. 268 p.
- Pyle, C. 1985. Vegetation disturbance history of Great Smoky Mountains National Park: An analysis of archival maps and records. U.S. Department of the Interior, National Park Service, Research/ Resources Management Report SER-77. 69 p.
- Pyle, C. 1988. The type and extent of anthropogenic vegetation disturbance in the Great Smoky Mountains before National Park Service acquisition. *Castanea* 53 (3): 25-235.
- Radford, A. E., Ahles, H. E. and C. R. Bell. 1968. *Manual of the vascular flora of the Carolinas*. University of North Carolina Press, Chapel Hill.



- Renfro, J. R. 1989. Evaluating the effects of ozone on the plants of Great Smoky Mountains National Park. *Park Science* 9 (4): 22-24.
- Rollins, R. C. 1993. *The Cruciferae of Continental North America: Systematics of the mustard family from the Arctic to Panama*. Stanford University Press, Stanford, California.
- Saunders, P. R. 1979. The vegetational impact of human disturbance on the spruce-fir forests of the southern Appalachian Mountains. Ph.D. dissertation, Duke University, Durham, North Carolina. 177 p.
- Schulz, O. E. 1903. Monographie der Gattung *Cardamine*. *Bot. Jahrb.* 32: 280-623.
- Small, J. K. 1933. *Manual of the southeastern flora*. University of North Carolina Press, Chapel Hill. 1554 p.
- Singer, F. J., W. T. Swank, and E. E. C. Clebsch. 1984. Effects of wild pig rooting in a deciduous forest. *J. Wildl. Manage.* 48(2): 464-473.
- Speers, C. F. 1958. The balsam woolly aphid in the southeast. *J. Forestry*. 56: 515-516.
- Strausbaugh, P. D. and E. L. Core. 1978. *Flora of West Virginia*, second ed. Seneca Books, Grantsville, WV.
- Weakley, A. S. 1996, draft. *Flora of the Carolinas and Virginia*. The Nature Conservancy, Southeast Regional Offices, Chapel Hill, North Carolina.
- White, P. S. 1982. The flora of Great Smoky Mountains National Park: an annotated checklist of the vascular plants and a review of previous floristic work. National Park Service, Great Smoky Mountains National Park, Research/Resource Management Report SER-55. 219 p.
- White, P. S. (ed.). 1984. *The southern Appalachian spruce-fir ecosystem: its biology and threats*. U. S. Department of the Interior, National Park Service, Research/Resources Management Report SER-71. 68 p.
- Wilds, S. P. 1996. Gradient analysis of the distribution of flowering dogwood (*Cornus florida* L.) and dogwood anthracnose (*Discula destructiva* Redlin.) in western Great Smoky Mountains National Park. M.S. Thesis, Univ. of North Carolina, Chapel Hill. 151p.
- Wiser, S. K. 1991. Two North Carolina locations for *Calamagrostis cainii* Hitch., previously considered endemic to Mt. Le Conte, Tennessee. *Castanea* 56 (2): 147-149.



- Wiser, S. K., R. K. Peet, and P. S. White. 1996. High-elevation rock outcrop vegetation of the Southern Appalachian Mountains. *Journal of Vegetation Science* 7: 703-722.
- Wofford, B. E. 1989. Guide to the vascular plants of the Blue Ridge. University of Georgia Press, Athens. 384 p.
- Wolock, D. M. 1993. Simulating the variable-source-area concept of streamflow generation with the watershed model TOPMODEL. U.S. Geol. Surv. Water Resour. Invest. Rep. 93-4124.
- Wolock, D. M. and G .J. McCabe, Jr. 1995. Comparison of single and multiple flow direction algorithms for computing topographic parameters in TOPMODEL. *Water Resources Research* 31(5): 1315-1324.



## APPENDICES





**Appendix A.** Herbaria visited during the course of this study.

Great Smoky Mountains National Park, Gatlinburg, Tennessee  
University of North Carolina, Chapel Hill, North Carolina  
University of Tennessee, Knoxville, Tennessee  
Western Carolina University, Cullowhee, North Carolina  
Clemson University, Clemson, South Carolina  
University of Georgia, Athens, Georgia

Specimens were requested and reviewed from the following institutions:

University of Alabama, Tuscaloosa, Alabama  
Vanderbilt University, Nashville, Tennessee  
Mars Hill College, Mars Hill, North Carolina



**Appendix B.** GIS variables used to model habitat in GSMNP.

<b>Name</b>	<b>Description <sup>a</sup></b>	<b>Value Range <sup>b</sup></b>	<b>Source</b>
Aspect	Aspect transformed using:	0 to 2.0	Beers et al. 1966
Elevation	Elevation (m)	914 - 2,027	USGS digital elevation
Landform Index	Index of meso-scale topographic	-4.026 to	McNab 1993
Planiform Curvature	Slope curvature in horizontal plane	-8.802 to 8.838	calculated from elevation with the CURVATURE command (Grid)
Profile Curvature	Slope curvature in vertical plane	-10.906 to 8.639	calculated from elevation with the CURVATURE command (Grid)
Relative Slope Position	Relative slope position (percent)	0 - 100	Wilds 1996
Rock <sup>c</sup>	Rock outcrops and landslide scars digitized from aerial photographs in April 1997	categorical (1 or 0)	digitized for this project
Slope	Slope steepness (degrees)	0.0 to 64.467	calculated from elevation with the CURVATURE command (Grid)
Solar Insolation	Index of exposure to sunlight; approximated for the solar equinox; considers both nearby and distant shadowing factors	1 to 227	calculated from elevation with the HILLSHADE command (Grid)
Shannon-Weiner Index	Shannon-Weiner index of topographic complexity	17 to 35	Miller 1986
Topographic Convergence Index	Simulates the flow accumulation of water; $TCI = \ln(A / \tan B)$ , where A is drained surface area and B is drained surface slope	17 to 182	Beven and Kirkby, 1979; Wolock 1993; Wolock and McCabe 1995; Halpin 1995



Topographic Relative Moisture Index	Index of moisture considering the effects of slope position, aspect and elevation	0 to 85	Parker 1982
Terrain Shape Index	Index of micro-scale topographic exposure	-169 to 71	McNab 1989
Vegetation	Vegetation type (9 forest types, heath bald, grassy bald, grape thicket, or treeless)	categorical (1 to 13)	MacKenzie 1993

<sup>a</sup> All variables were continuous variables except for vegetation and rock, which were nominal variables. Vegetation had a base resolution of 90 m; rock had a resolution of 15 m. All others had 30 m resolution.

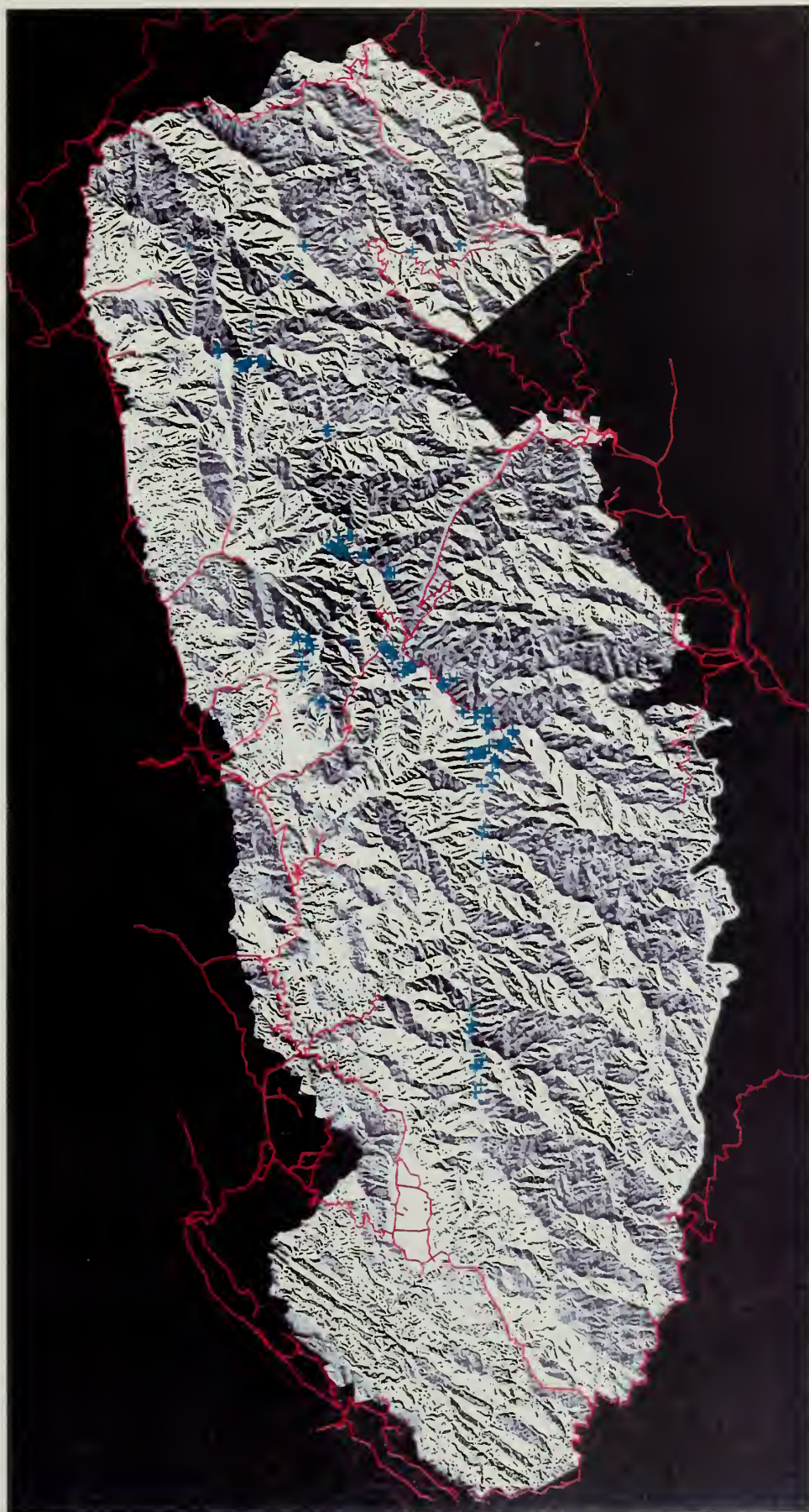
<sup>b</sup> Range values were for the study area, defined as GSMNP above 914 m.

<sup>c</sup> Only included in the model for *Calamagrostis cainii*.






# Known locations of *Cardamine clematitis* in Great Smoky Mountains NP



## Legend

 roads

 known positions



Scale = 1:380,000

0 25  
kilometers

Data sources: Location data collected by John R. Boetsch et al., 1996 - 1998.





# Known locations of *Rugelia nudicaulis* in Great Smoky Mountains NP



## Legend



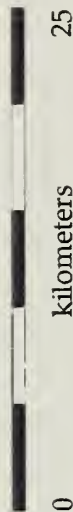
roads



known positions



Scale = 1:380,000

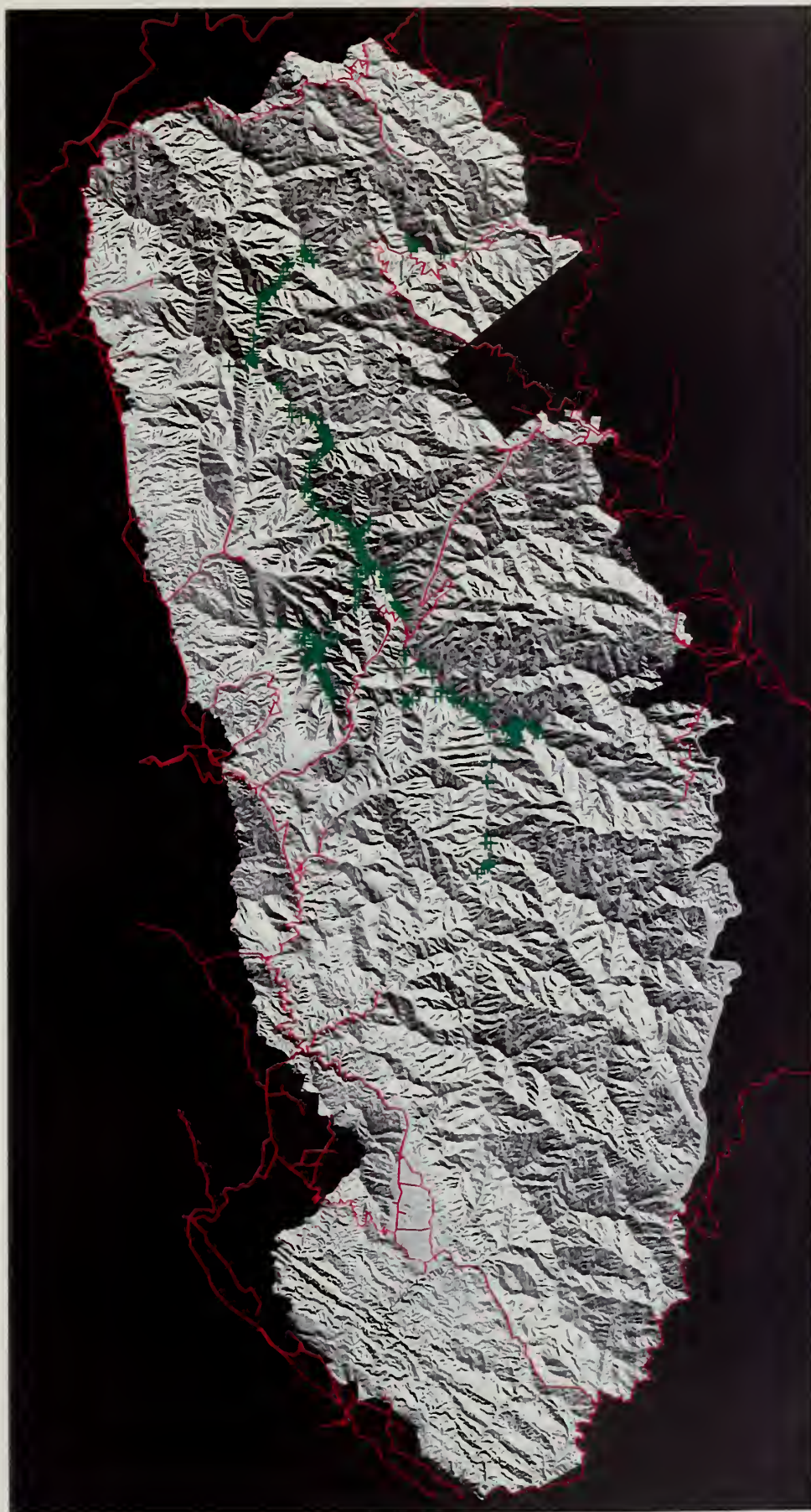


Data sources: Location data collected by John R. Boetsch et al., 1996 - 1998.

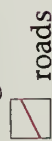




# Known locations of *Glyceria nubigena* in Great Smoky Mountains NP



## Legend



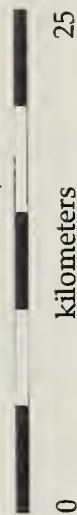
roads



known positions



Scale = 1:380,000

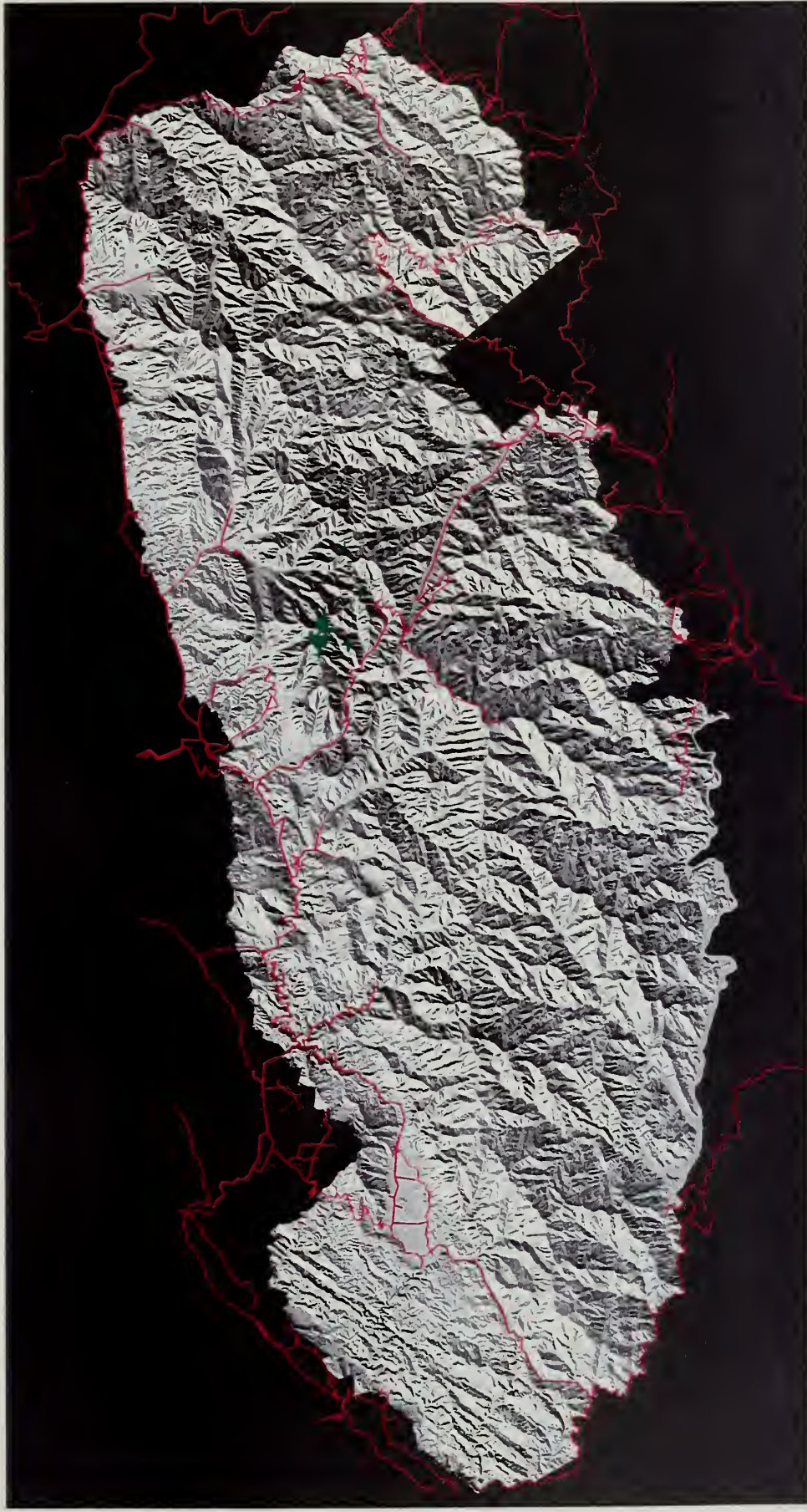


Data sources: Location data collected by John R. Boetsch et al., 1996 - 1998.





# Known locations of *Calamagrostis cainii* in Great Smoky Mountains NP



## Legend

 roads

 known positions



Scale = 1:380,000

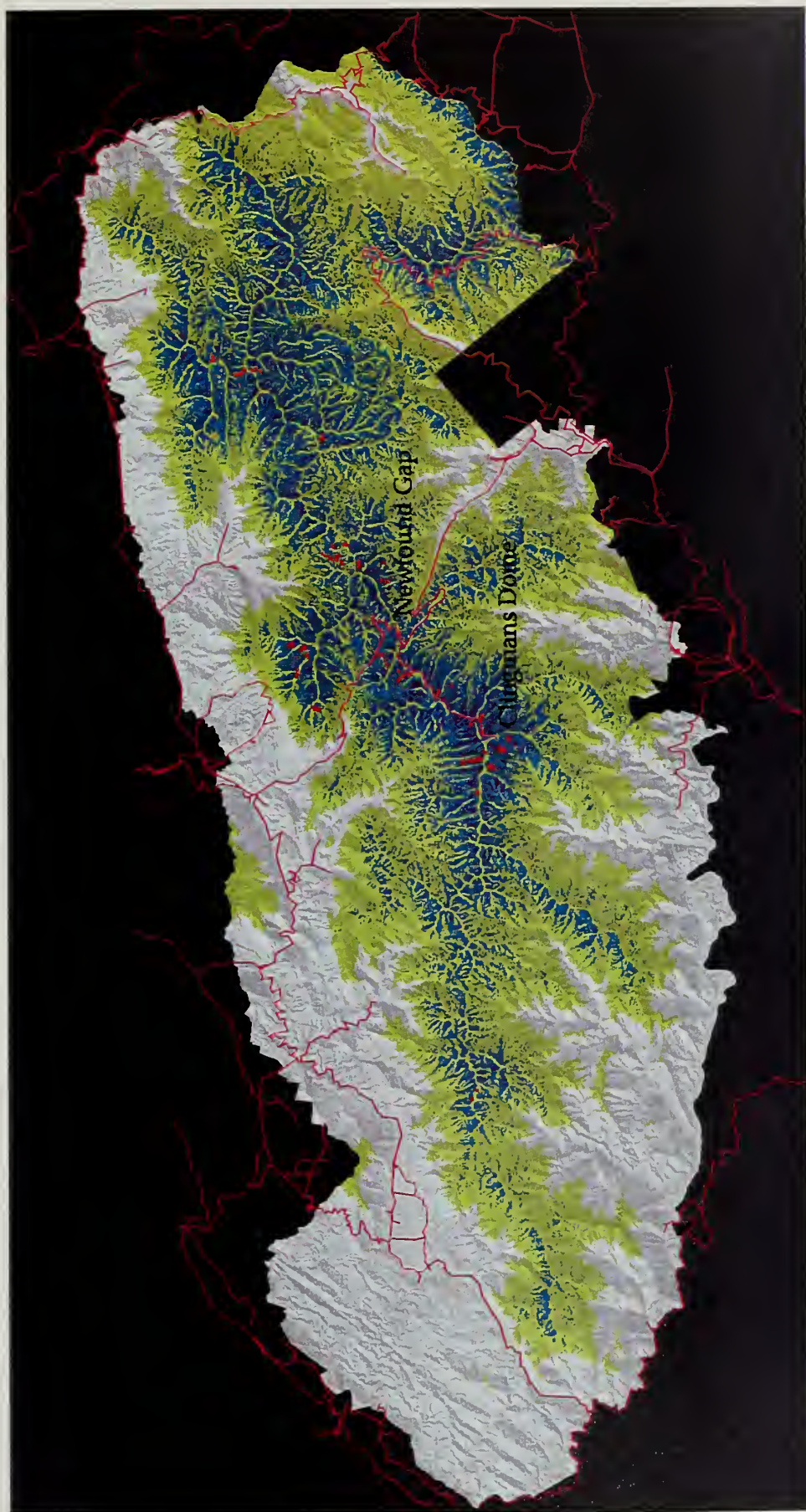


Data sources: Location data collected by John R. Boetsch et al., 1996 - 1998.





# Predicted distribution of *Cardamine clematitis* in Great Smoky Mountains NP

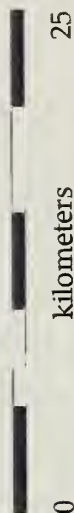


## Category

- Habitat most similar to input (50% of variation for input locations)
- Intermediate (95% of variation)
- Least similar habitat
- model input positions
- roads



Scale = 1:380,000

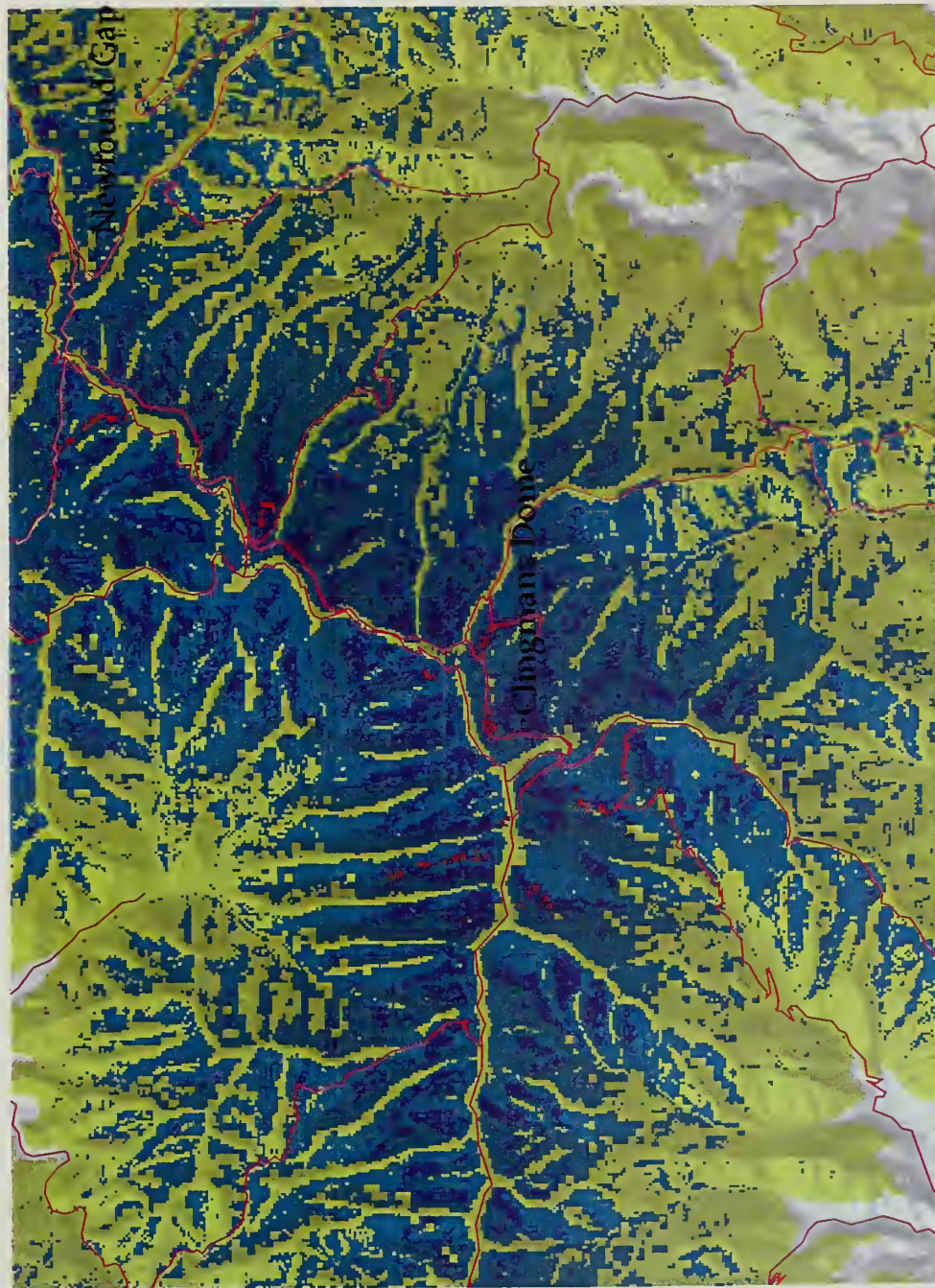


Data source: Mahalanobis distance model of habitat created by John R. Boetsch and Frank van Manen, August 1997.





# Predicted distribution of *Cardamine clematitis* in Clingmans Dome area, GSMNP

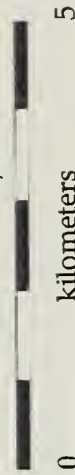


## Category

- Habitat most similar to input (50% of variation for input locations)
- Intermediate (95% of variation)
- Least similar habitat
- model input positions
- roads and trails



Scale = 1:85,000

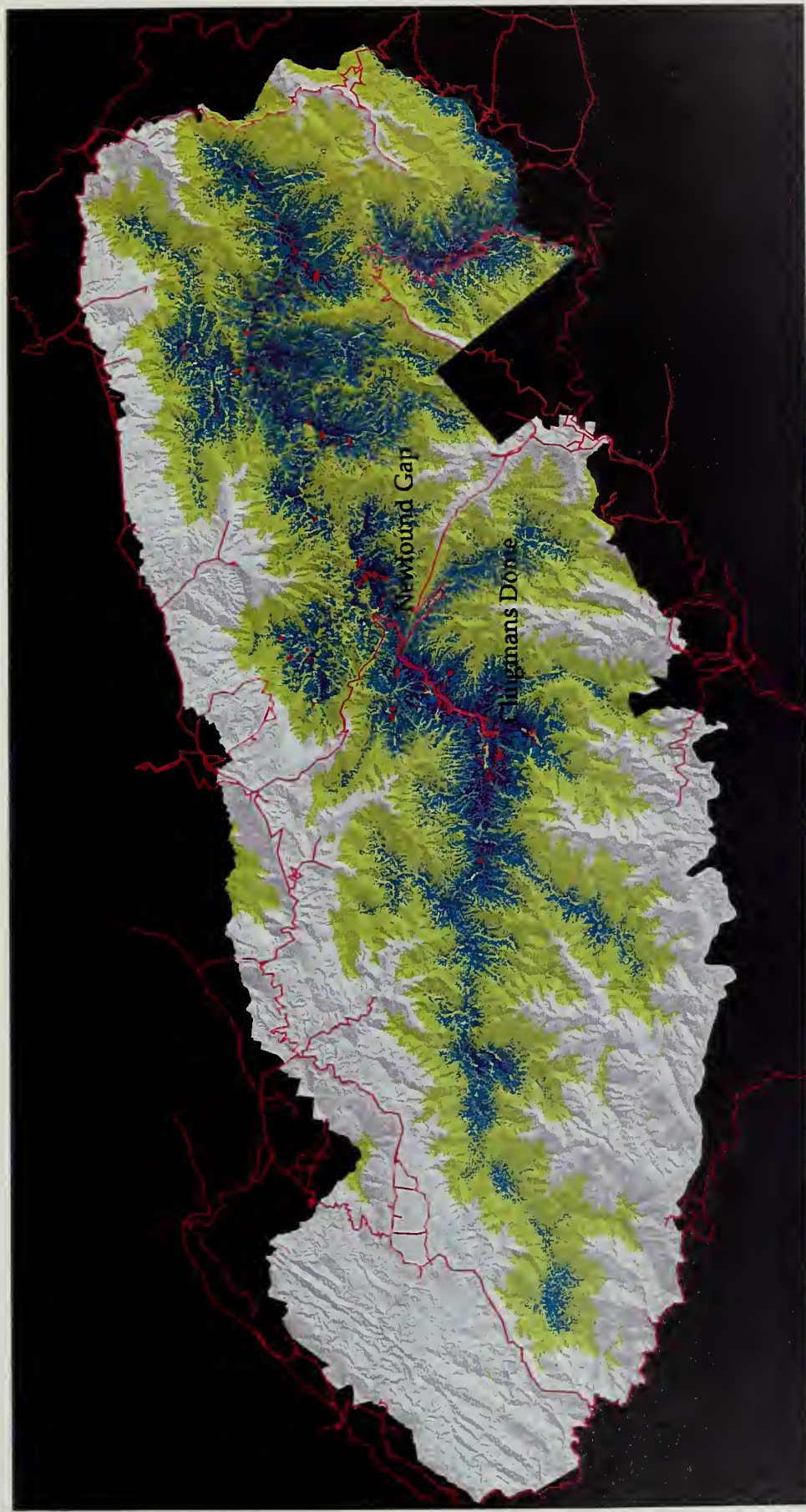


Data source: Mahalanobis distance model of habitat created by John R. Boetsch and Frank van Manen, August 1997.





# Predicted distribution of *Rugelia nudicaulis* in Great Smoky Mountains NP



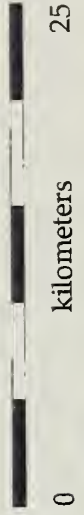
## Category

- Habitat most similar to input (50% of variation for input locations)
- Intermediate (95% of variation)
- Least similar habitat
- model input positions
- /

 roads



Scale = 1:380,000

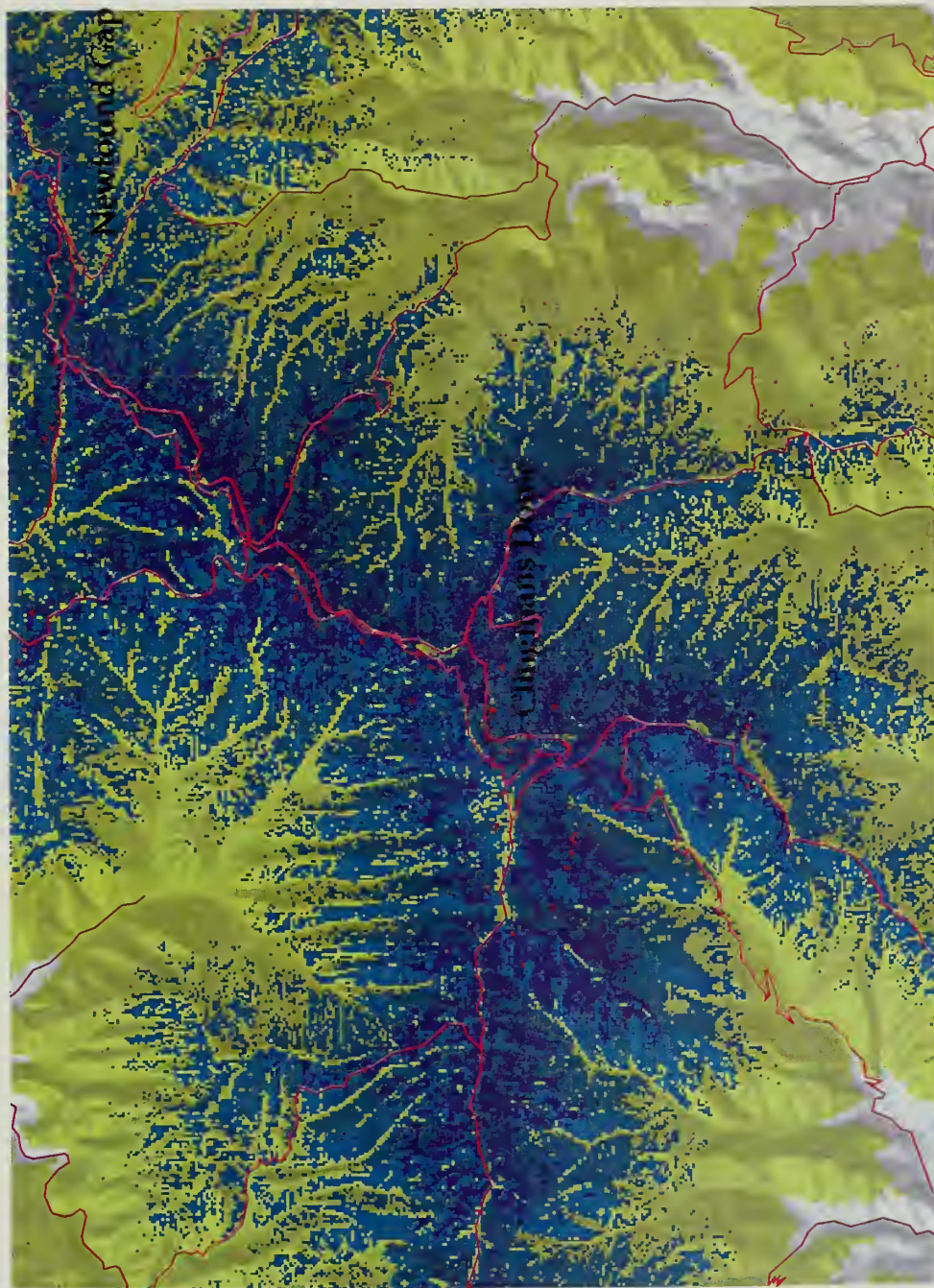


Data source: Mahalanobis distance model of habitat created by John R. Boetsch, September 1998.





# Predicted distribution of *Rugelia nudicaulis* in Clingmans Dome area, GSMNP



## Category

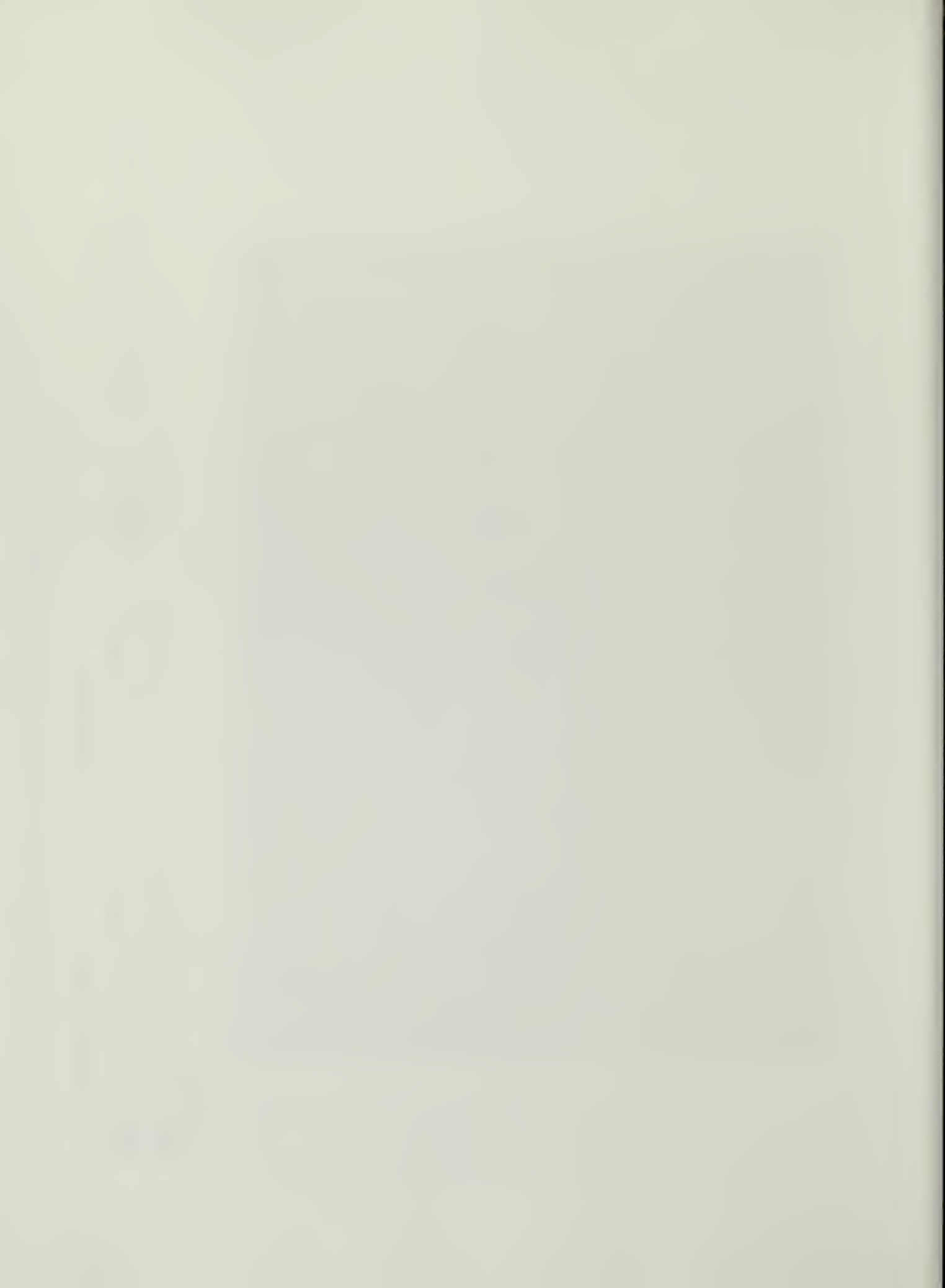
- Habitat most similar to input (50% of variation for input locations)
- Intermediate (95% of variation)
- Least similar habitat
- model input positions
- roads and trails



Scale = 1:85,000

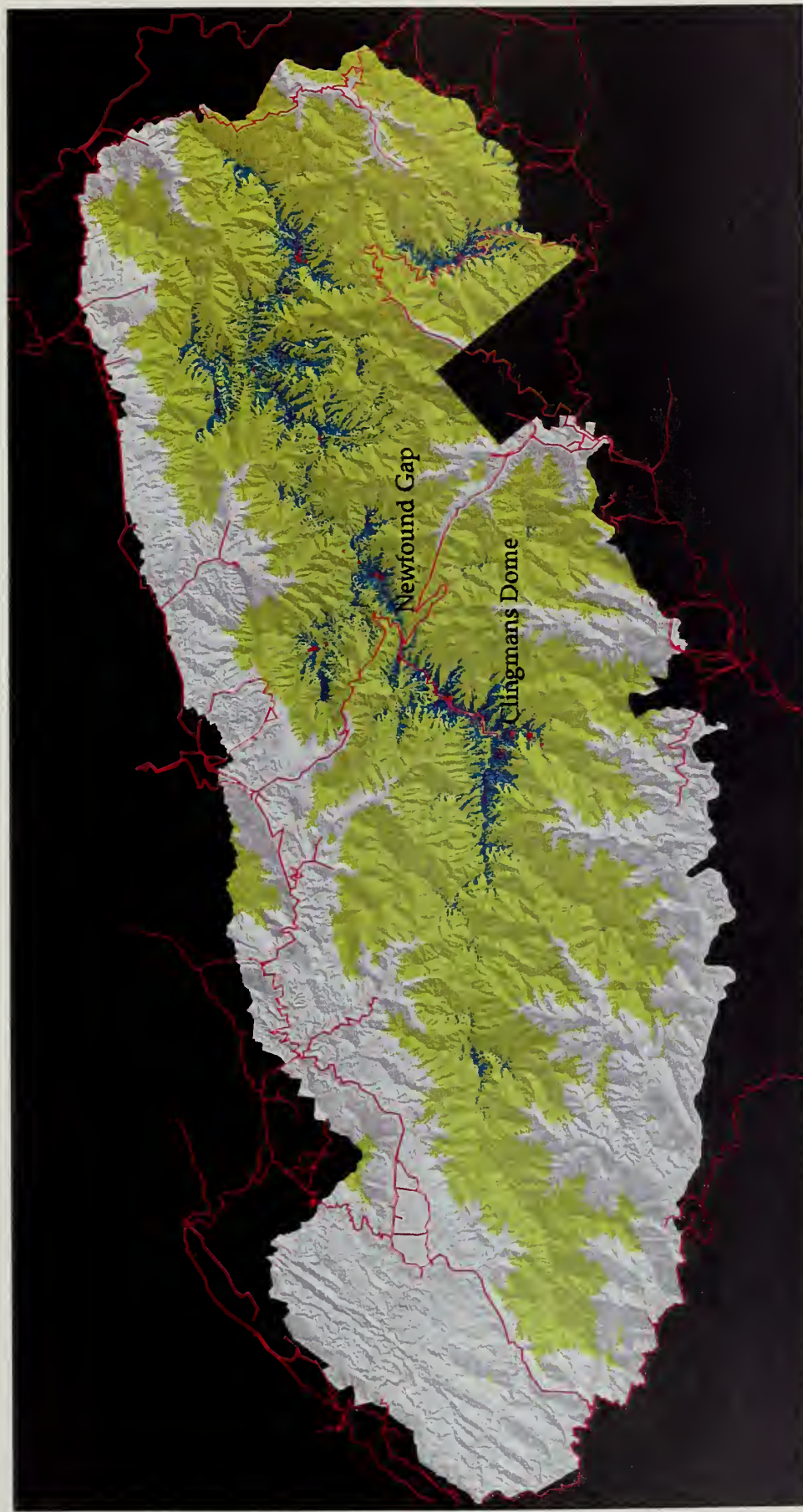
0 5 kilometers

Data source: Mahalanobis distance model of habitat created by John R. Boetsch, September 1998.





# Predicted distribution of *Glyceria nubigena* in Great Smoky Mountains NP

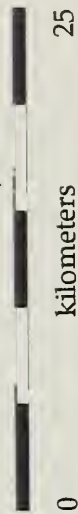


## Category

- Habitat most similar to input (50% of variation for input locations)
- Intermediate (95% of variation)
- Least similar habitat
- model input positions
- roads



Scale = 1:380,000

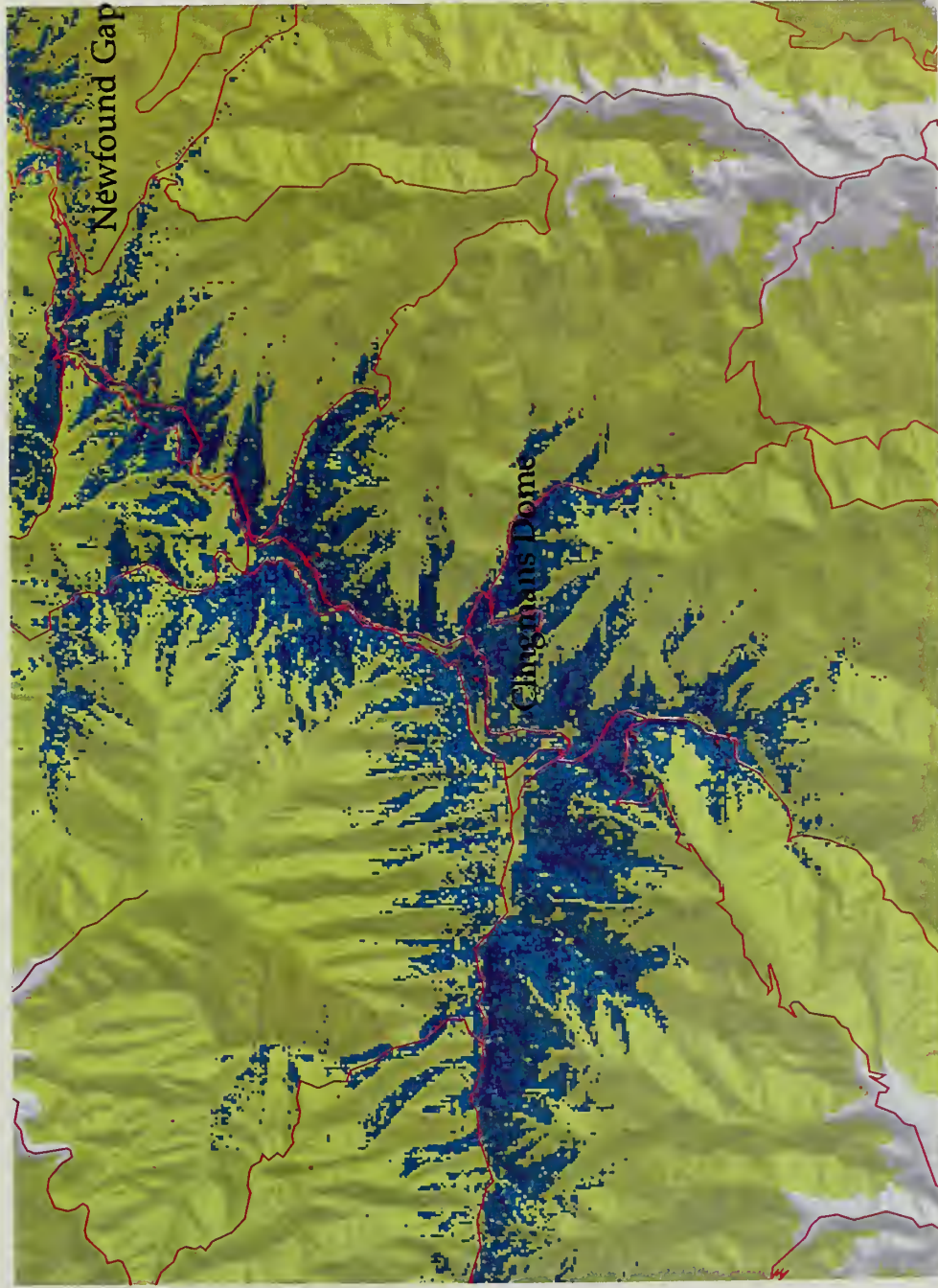


Data source: Mahalanobis distance model of habitat created by John R. Boetsch, September 1998.





# Predicted distribution of *Glyceria nubigena* in Clingmans Dome area, GSMNP



## Category

- Habitat most similar to input (50% of variation for input locations)
- Intermediate (95% of variation)
- Least similar habitat
- model input positions
- roads and trails



Scale = 1:85,000

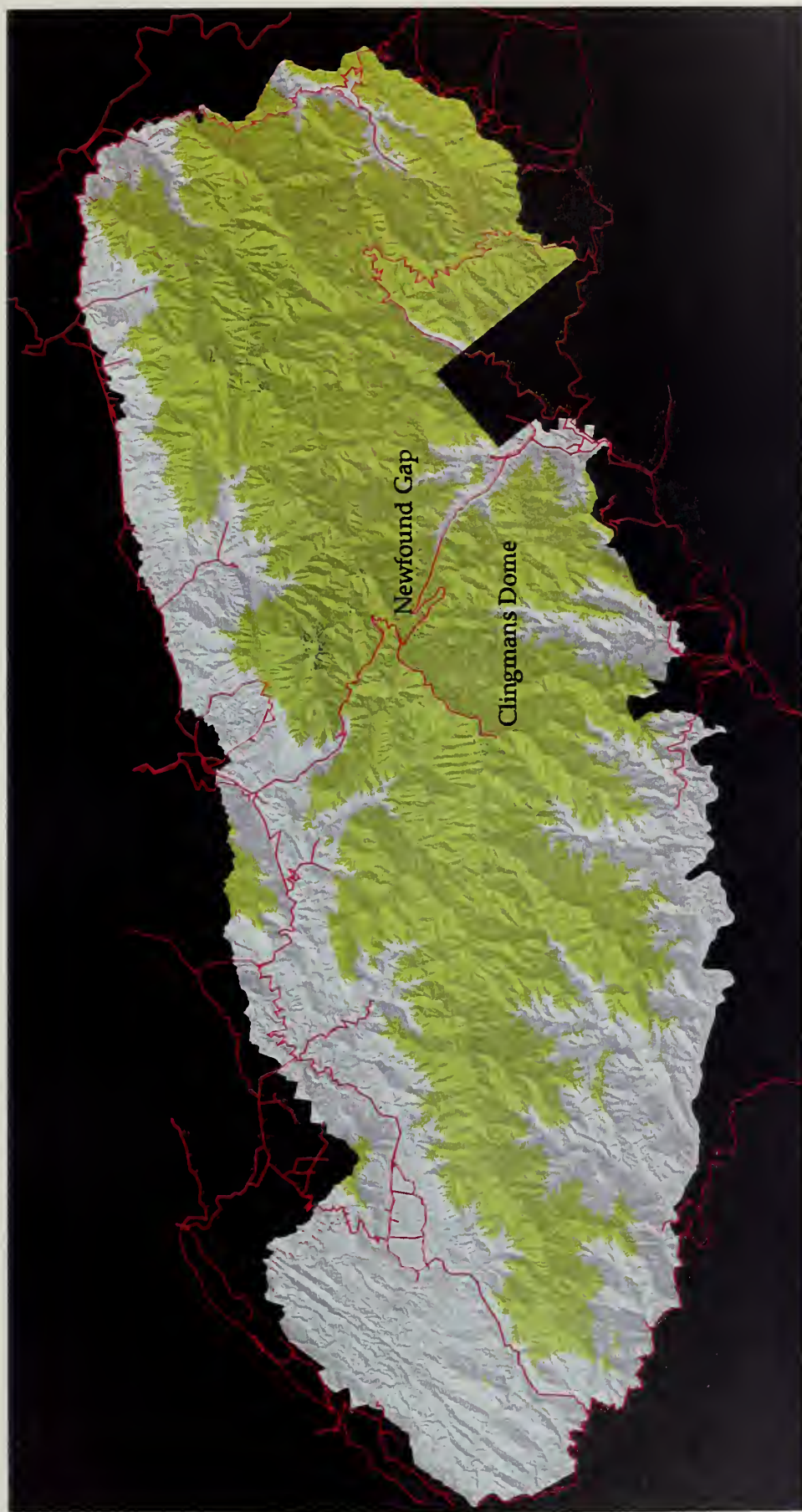


Data source: Mahalanobis distance model of habitat created by John R. Boetsch, September 1998.





# Predicted distribution of *Calamagrostis cainii* in Great Smoky Mountains NP

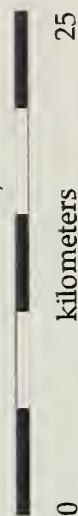


## Category

- Habitat most similar to input (50% of variation for input locations)
- Intermediate (95% of variation)
- Least similar habitat
- roads



Scale = 1:380,000

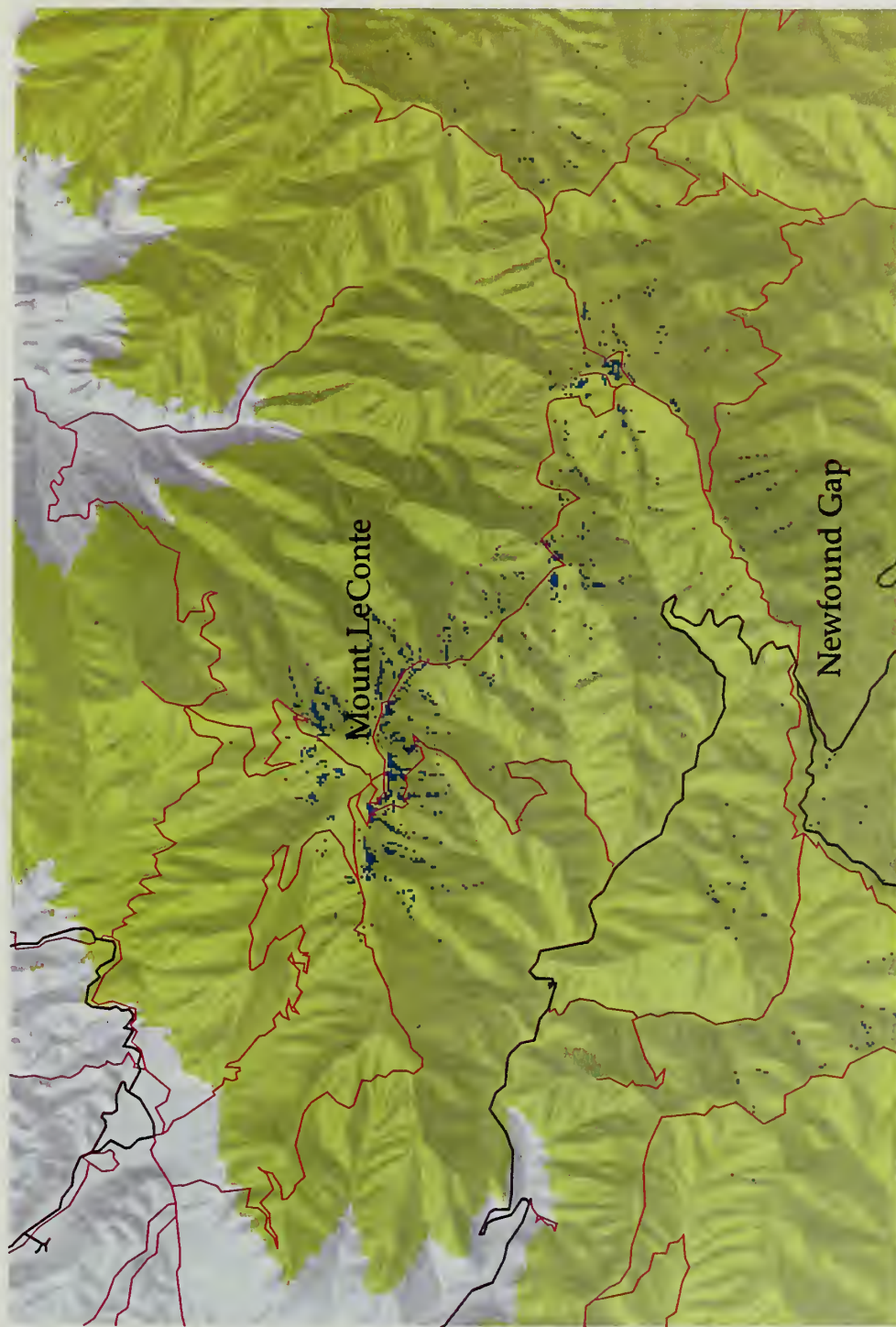


Data source: Mahalanobis distance model of habitat created by John R. Boetsch, September 1998.





# Predicted distribution of *Calamagrostis cainii* in Mount LeConte area, GSMNP



## Category

- Habitat most similar to input (50% of variation for input locations)
- Intermediate (95% of variation)
- Least similar habitat
- trails
- roads



Scale = 1:80,000

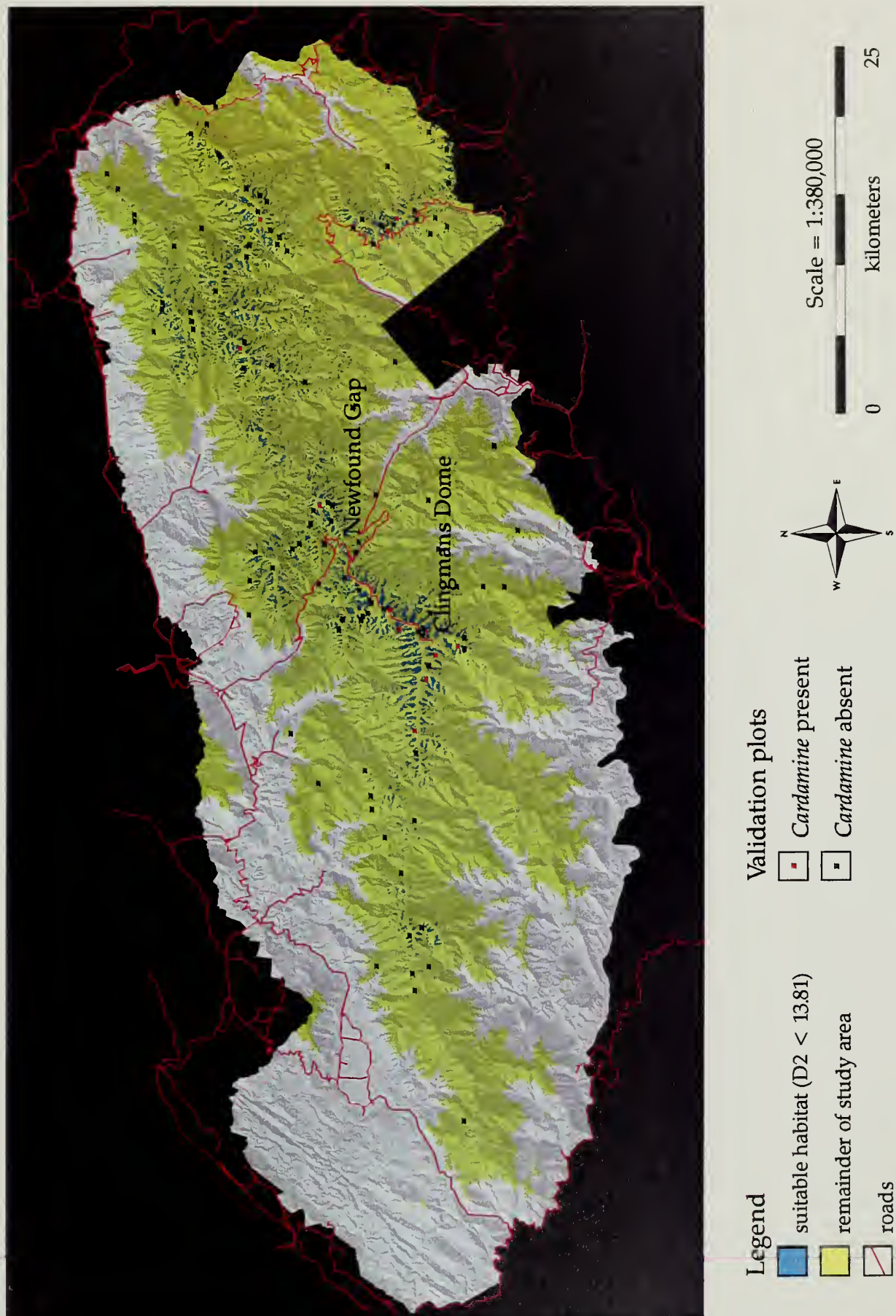


Data source: Mahalanobis distance model of habitat created by John R. Boetsch, September 1998.





# Predicted distribution of *Cardamine clematitis* in GSMNP, as defined by field validation

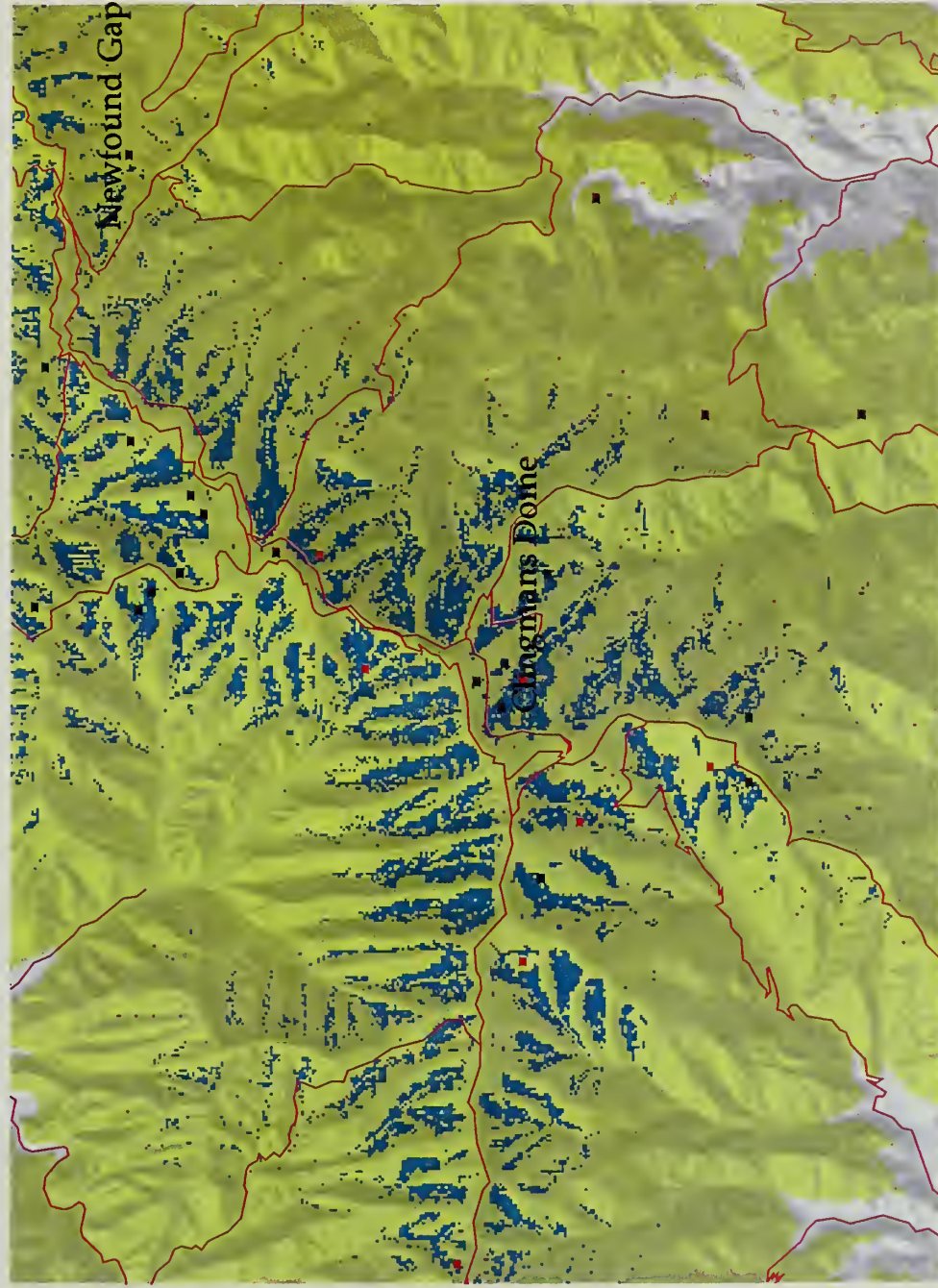


Data source: Mahalanobis distance model of habitat created by John R. Boetsch and Frank van Manen, August 1997.





# Validated distribution of *Cardamine clematitis* near Clingmans Dome, GSMNP



## Legend

- suitable habitat ( $D2 < 13.81$ )
- remainder of study area
- roads and trails

## Validation plots

- Cardamine* present
- Cardamine* absent



Scale = 1:85,000



Data source: Mahalanobis distance model of habitat created by John R. Boetsch and Frank van Manen, August 1997.





